The Open University

MST210

Mathematical methods, models and modelling

Book A

Cover image: This shows arrows representing a vector field and a corresponding contour map. The vector field shows the direction of change for a model of two competing populations of animals. You will meet this model in Unit 12.

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from the Student Registration and Enquiry Service, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)845 300 6090; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

To purchase a selection of Open University materials visit www.ouw.co.uk, or contact Open University Worldwide, Walton Hall, Milton Keynes MK7 6AA, United Kingdom for a brochure (tel. +44 (0)1908 858779; fax +44 (0)1908 858787; email ouw-customer-services@open.ac.uk).

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2014.

Copyright © 2014 The Open University

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, transmitted or utilised in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without written permission from the publisher or a licence from the Copyright Licensing Agency Ltd. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS (website www.cla.co.uk).

Open University materials may also be made available in electronic formats for use by students of the University. All rights, including copyright and related rights and database rights, in electronic materials and their contents are owned by or licensed to The Open University, or otherwise used by The Open University as permitted by applicable law.

In using electronic materials and their contents you agree that your use will be solely for the purposes of following an Open University course of study or otherwise as licensed by The Open University or its assigns.

Except as permitted above you undertake not to copy, store in any medium (including electronic storage or use in a website), distribute, transmit or retransmit, broadcast, modify or show in public such electronic materials in whole or in part without the prior written consent of The Open University or in accordance with the Copyright, Designs and Patents Act 1988.

Edited, designed and typeset by The Open University, using the Open University TEX System.

Printed in the United Kingdom by Bell & Bain Ltd, Glasgow.

ISBN 978 1 7800 7867 0

1.1

Contents

Unit 1 First- and second-order differential equations **1**

Introduction **3**

1 First-order differential equations **3**
1.1 Solutions of differential equations 6
1.2 Direct integration 11
1.3 Separation of variables 13
1.4 Integrating factor method 16
1.5 Direction fields and Euler's method 20

2 Homogeneous second-order differential equations **33**
2.1 Linear constant-coefficient differential equations 34
2.2 Method of solution 36

3 Inhomogeneous second-order differential equations **44**
3.1 General method of solution 45
3.2 Finding a particular integral by the method of undetermined coefficients 48
3.3 Exceptional cases 54
3.4 Combining cases 57

4 Initial conditions and boundary conditions **59**
4.1 Initial-value problems 59
4.2 Boundary-value problems 61

Learning outcomes **65**

Solutions to exercises **67**

Acknowledgements **91**

Unit 2 Vector algebra and statics **93**

Introduction **95**

1 Describing, representing and combining vectors **97**
1.1 Scalars and vectors 97
1.2 Using arrows to represent vectors 98
1.3 Vector algebra 99

2 Cartesian components and products of vectors **104**
2.1 Vectors in three dimensions 104
2.2 The dot product 111
2.3 The cross product 118

3 Modelling forces **123**
3.1 Particles 123
3.2 Weight 124
3.3 Normal reaction 126
3.4 Tension 126
3.5 Friction 129
3.6 Two-particle systems 135

4 Torques **141**
4.1 Extended and rigid bodies 141
4.2 Turning effect of a force 141

5 Applying the principles **145**

Learning outcomes **150**

Solutions to exercises **151**

Unit 3 Dynamics **175**

Introduction **177**

1 Describing motion **179**
1.1 Motion of a particle 179
1.2 One-dimensional motion 185

2 A theory of motion **191**

3 Predicting motion **195**

4 Some more force models **202**
4.1 Friction 202
4.2 Air resistance 205

5 Projectiles **215**
5.1 Motion of a projectile 215
5.2 Trajectory of a projectile 221

Learning outcomes **228**

Solutions to exercises **229**

Index **255**

Unit 1

First- and second-order differential equations

Introduction

An important class of equations that arise in mathematics consists of those that feature the *rates of change* of one or more variables with respect to one or more others. These rates of change are expressed mathematically by *derivatives*, and the corresponding equations are called *differential equations*. Equations of this type crop up in a wide variety of situations. They are found, for example, in models of physical, electronic, economic, demographic and biological phenomena.

First-order differential equations, which are the particular topic of Section 1, feature derivatives of order one only; that is, if the rate of change of variable y with respect to variable x is involved, then the equations feature dy/dx but not d^2y/dx^2, d^3y/dx^3, etc.

When a differential equation arises, it is usually an important aim to *solve* the equation. For an equation that features the derivative dy/dx, this entails expressing the *dependent variable* y directly in terms of the *independent variable* x. You will see four possible approaches to finding a solution in Section 1.

This unit also considers *second-order* differential equations, that is, differential equations that involve a second (but no higher) derivative. Examples of second-order differential equations are

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 4e^x \quad \text{and} \quad 3\frac{d^2y}{dx^2} + y^2 \sin x = x^2.$$

A second-order differential equation may or may not include a first derivative.

1 First-order differential equations

The content of this section should be familiar to you if you have studied other modules in which first-order differential equations are discussed. For the most part, this section looks at methods for finding *analytic* solutions of first-order differential equations – that is, solutions expressed in terms of exact formulas in the three possible cases

$$\frac{dy}{dx} = f(x), \quad \frac{dy}{dx} = g(x)\,h(y), \quad \frac{dy}{dx} + g(x)\,y = h(x),$$

where f, g and h are specified functions in each case. The first equation can be solved by *direct integration*, the second by *separation of variables*, and the third by finding an *integrating factor*. For other types of equation, and indeed for some of the equations of this form where the integration cannot be performed analytically, numerical methods can be used to find an approximate solution. However, where a simple formula can be found, this is likely to be more informative than a solution found by the use of a numerical method.

In this module you will meet many examples of differential equations. Frequently these arise from studying the motion of physical objects, but we start with an example drawn from biology and show how this leads naturally to a particular differential equation.

Suppose that we are interested in the size of a particular population, and in how it varies over time. The first point to make is that any population size is measured in integers (whole numbers), so it is not clear how differentiation will be relevant. (Differentiable functions must be continuous, therefore defined on an interval of real numbers in $\mathbb{R}$.) Nevertheless, if the population is large, say in hundreds of thousands, a change of one unit will be relatively very small, and in these circumstances we may choose to model the population size as a *continuous* function of time. We write this function as $P(t)$, and our task is to show how $P(t)$ may be described by a differential equation.

Let us assume a fixed starting time (which we label $t = 0$). If the population is not constant, then there will be 'leavers' and 'joiners'. For example, in a population of humans in a particular country, the former will be those who die or emigrate, while the latter represent births and immigrants.

It is usual to express birth rates as a proportion of the current population size. Death rates are specified in a similar way. To emphasise that these rates are expressed as a *proportion* of the current population, we use the terms **proportionate birth rate** and **proportionate death rate**.

Note that in our model the proportionate birth rate is expressed as a proportion of the whole population, not just the number of women.

For our simple model we ignore immigration and emigration, and concentrate solely on births and deaths. Denote the proportionate birth rate by b and the proportionate death rate by c. Then in a short interval of time δt, we would expect

$$\text{number of births} \simeq b\,P(t)\,\delta t, \tag{1}$$

$$\text{number of deaths} \simeq c\,P(t)\,\delta t, \tag{2}$$

where $P(t)$ is the population size (in units of 100 000) at time t.

At this stage, we seek some relationship between the chosen variables. In order to find this, we make use of the **input–output principle**, which can be expressed as

$$\boxed{\text{accumulation}} = \boxed{\text{input}} - \boxed{\text{output}}.$$

This principle applies to any quantity whose change, over a given time interval, is due solely to the specified input and output.

The *accumulation* δP of population over the time interval δt is the population at the end of the interval minus the population at the start of the interval, that is,

$$\delta P = P(t + \delta t) - P(t).$$

The *input* is the number of births (equation (1)), and the *output* is the number of deaths (equation (2)). The input–output principle now enables us to express the accumulation δP of the population over the time

interval δt as

$$\delta P \simeq b\,P(t)\,\delta t - c\,P(t)\,\delta t = (b-c)\,P(t)\,\delta t.$$

Dividing through by δt, we obtain

$$\frac{\delta P}{\delta t} \simeq (b-c)\,P(t).$$

The approximations involved in deriving this equation become progressively more accurate for shorter time intervals. So, finally, by letting δt tend to zero, we obtain

This is the step that requires P to be a continuous (rather than discrete) function of t.

$$\frac{dP}{dt} = (b-c)\,P(t).$$

(This follows because

$$\frac{dP}{dt} = \lim_{\delta t \to 0} \frac{P(t+\delta t) - P(t)}{\delta t}$$

is the *definition* of the derivative of P.)

This is a *differential* equation because it describes dP/dt rather than the eventual object of our interest (which is P itself). The purpose of this unit is to enable you to solve a wide variety of such equations.

We can simplify the above equation slightly by using the **proportionate growth rate** r, which is the difference between the proportionate birth and death rates: $r = b - c$. Then our model becomes

$$\frac{dP}{dt} = rP. \tag{3}$$

For very simple population models, r is taken to be a constant. As we will see, this leads to a prediction of exponential growth (or, if $r < 0$, decay) in population size with time, as illustrated in Figure 1. This may be a very good approximation for certain populations, but it cannot be sustained indefinitely if $r > 0$.

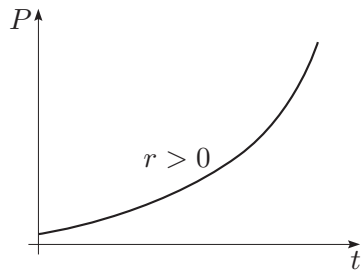


Figure 1 Population growth

In practice, both the proportionate birth rate and the proportionate death rate will vary, and so therefore will the proportionate growth rate. It turns out to be convenient to model these changes as being dependent on the population size, so the proportionate growth rate r becomes a function of P. The justification for this is as follows. When the population is low, one may assume that there is potential for it to grow (assuming a reasonable environment). The proportionate growth rate should therefore be high. However, as the population grows, there will be competition for resources. Thus the proportionate growth rate will decline, and in this way unlimited (exponential) growth does not occur.

A particularly useful model arises from taking $r(P)$ to be a decreasing linear function of P. We write this as

$$r(P) = k\left(1 - \frac{P}{M}\right), \tag{4}$$

You will see later why this particular form is chosen.

where k and M are positive constants. Looking at this formula, you can see that the proportionate growth rate r decreases linearly with P, from the value k (when $P = 0$) to the value 0 (when $P = M$).

Using this expression for r, the differential equation (3) satisfied by P becomes

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right). \tag{5}$$

This is well known to biologists as the **logistic equation** – we will consider it further in Subsection 1.1, and see how to solve it in Subsection 1.3. For now, we have achieved our objective of showing that differential equations arise naturally in modelling the real world.

Exercise 1

Suppose that a population obeys the logistic model (with the proportionate growth rate given by equation (4)), and that you are given the following information. When $P = 10$ the proportionate growth rate is 1, and when $P = 10\,000$ the proportionate growth rate is 0.

Find the corresponding values of k and M.

1.1 Solutions of differential equations

This notation is named after Gottfried Leibniz (1646–1716).

The *derivative* of a variable y with respect to another variable x is denoted in *Leibniz notation* by dy/dx. In this derivative expression we refer to y as the *dependent variable* and to x as the *independent variable*.

There are other notations in use for derivatives. If the relation between variables x and y is expressed in terms of a function f, so that $y = f(x)$, then the derivative may be written in function notation as $f'(x)$.

A further notation, attributed to Sir Isaac Newton, is restricted to cases in which the independent variable is time, denoted by t. The derivative of $y = f(t)$ could be written in this case as $\dot{y}$, in which the dot over the y stands for the d/dt of Leibniz notation. Thus we may express this derivative in any of the equivalent forms

$$\frac{dy}{dt} = \dot{y} = f'(t).$$

Further derivatives are obtained on differentiating this first derivative. The second derivative of $y = f(t)$ could be represented by any of the forms

$$\frac{d^2y}{dt^2} = \ddot{y} = f''(t).$$

These possible notations have different strengths and weaknesses, and which is most appropriate in any situation depends on the purpose at hand. You will see all of these notations employed at various times during the module.

It is common practice in applied mathematics to reduce the proliferation of symbols as far as possible. One aspect of this practice is that we often avoid allocating separate symbols to variables and to associated functions. Thus in place of the equation $y = f(t)$ (where y and t denote variables,

and f denotes the function that relates them), we could write $y = y(t)$, which is read as 'y is a function of t'.

Strictly speaking, this is an abuse of notation, since there is ambiguity as to exactly what the symbol y represents: it is a variable on the left-hand side of $y = y(t)$, but a function on the right-hand side. However, it is a very convenient abuse.

The following definitions explain just what are meant by a *differential equation*, by the *order* of such an equation, and by a *solution* of it.

A **differential equation** for $y = y(x)$ is an equation that relates the independent variable x, the dependent variable y, and one or more derivatives of y.

The **order** of such a differential equation is the order of the highest derivative that appears in the equation. Thus a **first-order** differential equation for $y = y(x)$ features only the first derivative, dy/dx.

A **solution** of such a differential equation is a function $y = y(x)$ that satisfies the differential equation.

A function must *satisfy* a differential equation in order to be regarded as a solution of it. The differential equation is satisfied by the function provided that when the function is substituted into the equation, the left- and right-hand sides of the equation give an identical expression. This substitution includes the requirement that the function should be *differentiable* (i.e. that it should have a derivative) at all points where it is claimed to be a solution.

You are asked to verify in the next exercise that three functions are solutions of corresponding first-order differential equations.

Exercise 2

Verify that each of the functions given below is a solution of the corresponding differential equation.

(a) $y = 2e^x - (x^2 + 2x + 2)$; $\dfrac{dy}{dx} = y + x^2$.

(b) $y = \tan x + \sec x$; $\dfrac{dy}{dx} = y \tan x + 1$.

(c) $y = t + Ce^{-t}$; $\dot{y} = -y + t + 1$. (Here C is an arbitrary constant.)

In the last part of Exercise 2 you were asked to verify that

$$y = t + Ce^{-t}$$

is a solution of the differential equation $\dot{y} = -y + t + 1$, where C is an arbitrary constant. In saying that C is *arbitrary*, we mean that it can assume any real value. Whatever number is chosen for C, the corresponding expression for $y(t)$ is always a solution of the differential equation. Choosing $C = 1$, for example, gives the particular function $y = t + e^{-t}$.

This demonstrates that solutions of a differential equation can exist in profusion; as a result, we need terms to distinguish between the totality of all these solutions for a given equation and the individual solutions that are completely specified.

The **general solution** of a differential equation is the collection of all possible solutions of that equation.

A **particular solution** of a differential equation is a single solution of the equation, and consists of a solution function whose rule contains no arbitrary constant.

In many cases it is possible to describe the general solution of a first-order differential equation by a single formula involving an arbitrary constant. For example, you will see from Exercise 2(c) that $y = t + Ce^{-t}$ is the general solution of the equation $\dot{y} = -y + t + 1$; this means that not only is $y = t + Ce^{-t}$ a solution whatever the value of C, but also *every* particular solution of the equation may be obtained by giving C a suitable value.

Exercise 3

(a) Verify that for any value of the constant C, the function $y = C - \frac{1}{3}e^{-3x}$ is a solution of the differential equation

$$\frac{dy}{dx} = e^{-3x}.$$

(b) Verify that for any value of the constant C, the function

$$P = \frac{CMe^{kt}}{1 + Ce^{kt}}$$

is a solution of the logistic equation

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right).$$

See equation (5).

As you have seen, there are many solutions of a differential equation. However, a particular solution of the equation, representing a definite relationship between the variables involved, is often what is needed. This is achieved by using a further piece of information in addition to the differential equation. Often the extra information takes the form of a pair of values for the independent and dependent variables.

For example, in the case of a population model, it would be natural to specify the starting population, P_0 say, and to start measuring time from $t = 0$. We could then write

$$P = P_0 \text{ when } t = 0, \quad \text{or equivalently,} \quad P(0) = P_0.$$

A requirement of this type is called an *initial condition*.

An **initial condition** associated with the differential equation

$$\frac{dy}{dx} = f(x, y)$$

specifies that the dependent variable y takes some value y_0 when the independent variable x takes some value x_0. This is written either as

$$y = y_0 \text{ when } x = x_0$$

or as

$$y(x_0) = y_0.$$

The numbers x_0 and y_0 are referred to as **initial values**.

The combination of a first-order differential equation and an initial condition is called an **initial-value problem**.

The word 'initial' in these definitions arises from those (frequent) cases in which the independent variable represents time. In such cases, the differential equation describes how the system being modelled behaves once started, while the initial condition specifies the configuration with which the system is started off. In fact, if the initial condition is $y(x_0) = y_0$, then we are often interested in solving the corresponding initial-value problem for $x > x_0$.

If x represents time, then $x > x_0$ is 'the future' after the system has been started off.

We usually require that an initial-value problem should have a *unique* solution, since then the outcome is completely determined by how the system behaves and its configuration at the start. Almost all the differential equations in this module do have unique solutions, and we will assume that all the initial-value problems in this unit have unique solutions.

Example 1

Using the result given in Exercise 3(a), solve the initial-value problem

$$\frac{dy}{dx} = e^{-3x}, \quad y(0) = 1.$$

Solution

From Exercise 3(a), a solution of the differential equation is

$$y = C - \tfrac{1}{3}e^{-3x},$$

In fact, as will be shown in Example 2, this is the general solution.

where C is a constant.

The initial condition says that $y = 1$ when $x = 0$, and on feeding these values into the above solution we find that

$$1 = C - \tfrac{1}{3}.$$

Hence $C = \frac{4}{3}$, and the particular solution of the differential equation that solves the initial-value problem is

$$y = \tfrac{4}{3} - \tfrac{1}{3}e^{-3x}.$$

Exercise 4

The size of a population (measured in units of hundreds of thousands) is modelled by the logistic equation

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right), \quad P(0) = 1,$$

where $k = 0.15$ and $M = 10$.

(a) Use your answer to Exercise 3(b) to solve this initial-value problem.

(b) Can you predict the long-term behaviour of the population size from your answer?

Finally in this subsection, note that one needs to keep an eye on the *domain* of the function defining the differential equation. 'Gaps' in the domain usually show up as some form of restriction on the nature of a solution curve. For example, consider the differential equation

$$\frac{dy}{dx} = \frac{1}{x}. \tag{6}$$

Since $|x| = -x$ if $x < 0$, we can write

$$\int \frac{1}{x}\,dx = \ln|x|.$$

It turns out that there are two distinct families of solutions of this equation, given by $y = \ln x + C$ (if $x > 0$) and $y = \ln(-x) + C$ (if $x < 0$). These two families of solutions are illustrated in Figure 2. Notice that the right-hand side of equation (6) is not defined at $x = 0$, and that there is no solution that crosses the y-axis.

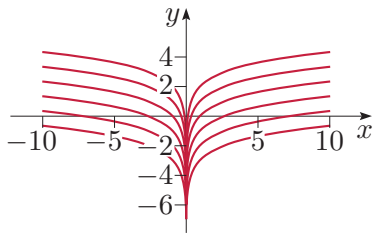


Figure 2 Solutions of equation (6)

1.2 Direct integration

Equations of the form

$$\frac{dy}{dx} = f(x)$$

can be solved by taking the indefinite integral of both sides of the equation.

The function $f(x)$ is assumed to be continuous (i.e. its graph has no breaks).

Example 2

(a) Find the general solution of the differential equation

$$\frac{dy}{dx} = e^{-3x}.$$

(b) Find the particular solution of this differential equation that satisfies the initial condition $y(0) = \frac{5}{3}$.

Solution

(a) On applying direct integration, we obtain the general solution

$$y = \int e^{-3x}\,dx = -\tfrac{1}{3}e^{-3x} + C,$$

where C is an arbitrary constant.

(b) In order to satisfy the initial condition $y(0) = \frac{5}{3}$ (i.e. $y = \frac{5}{3}$ when $x = 0$), we must have

$$\tfrac{5}{3} = -\tfrac{1}{3}e^{0} + C,$$

so $C = 2$. The required particular solution is therefore

$$y = -\tfrac{1}{3}e^{-3x} + 2.$$

Procedure 1 Direct integration

The general solution of the differential equation

$$\frac{dy}{dx} = f(x)$$

is

$$y = \int f(x)\,dx = F(x) + C, \tag{7}$$

where $F(x)$ is an integral of $f(x)$, and C is an arbitrary constant.

Once the general solution has been found, it is possible to single out a particular solution by specifying a value for the constant C. This value may be found by applying an initial condition.

Exercise 5

Solve each of the following initial-value problems.

(a) $\dfrac{dy}{dx} = 6x, \quad y(1) = 5.$

(b) $\dfrac{dv}{du} = e^{4u}, \quad v(0) = 2.$

(c) $\dot{y} = \dfrac{t}{1+t^2}, \quad y(0) = 2.$

(*Hint*: For the integral, try the substitution $u = 1 + t^2$.)

The answer to Exercise 5(c) can be generalised.

Any differential equation of the form

$$\frac{dy}{dx} = k\,\frac{f'(x)}{f(x)} \quad (f(x) \neq 0),$$

where k is a constant, can be integrated to give the general solution

$$y = k \ln|f(x)| + C, \tag{8}$$

where C is an arbitrary constant.

The proof of this result involves differentiating the solution and showing that y satisfies the differential equation.

In Exercise 5(c) we had $f(t) = 1 + t^2$ and $f'(t) = 2t$, with $k = \frac{1}{2}$. The initial right-hand side $t/(1 + t^2)$ had to be manipulated slightly to get it into the right form. Spotting integrands of this form (or of this form apart from a constant multiple) can allow you to solve some quite tricky-looking problems.

Exercise 6

Find the general solution of each of the following differential equations, where a is a non-zero constant.

(a) $\dfrac{dy}{du} = \dfrac{1}{u-a} \quad (u \neq a)$

(b) $\dfrac{dy}{dx} = \dfrac{1}{x(1-ax)} \quad (x \neq 0,\ x \neq 1/a)$

$\left(\textit{Hint}\text{: First verify that } \dfrac{1}{x(1-ax)} = \dfrac{1}{x} + \dfrac{a}{1-ax}.\right)$

1.3 Separation of variables

Equations of the form

$$\frac{dy}{dx} = g(x)\,h(y),$$

where g is a function of x only, and h is a function of y only, can be solved by dividing both sides by $h(y)$ to give

$$\frac{1}{h(y)}\frac{dy}{dx} = g(x),$$

and then taking the indefinite integral of both sides of the equation, as before. The final step is to rearrange the results to obtain the explicit solution in the form

$$y = \text{a function of } x,$$

if this is possible. It is necessary to be careful about the domain or image set of the solution obtained, as the following example illustrates.

The functions $g(x)$ and $h(y)$ are assumed to be continuous, and $h(y) \neq 0$.

Example 3

(a) Find the general solution of the differential equation

$$\frac{dy}{dx} = \frac{3x^2y}{4+x^3} \quad (y > 0).$$

(b) Find the general solution of the differential equation

$$\frac{dy}{dx} = -\frac{x}{3y} \quad (y > 0),$$

and the particular solution that satisfies the initial condition $y(0) = 3$.

Solution

(a) The equation is of the form

$$\frac{dy}{dx} = g(x)\,h(y),$$

where the obvious choices for g and h are

$$g(x) = \frac{3x^2}{4+x^3} \quad \text{and} \quad h(y) = y.$$

Dividing both sides of the given differential equation by y gives

$$\frac{1}{y}\frac{dy}{dx} = \frac{3x^2}{4+x^3}.$$

Integrating both sides with respect to x gives

$$\int \frac{1}{y}\frac{dy}{dx}\,dx = \int \frac{1}{y}\,dy = \int \frac{3x^2}{4+x^3}\,dx,$$

where the second integral comes from applying the rule for integration by substitution.

Spotting that the right-hand integrand is of the form $f'(x)/f(x)$ with $f(x) = 4 + x^3$, we have

$$\ln|y| = \ln|4 + x^3| + C,$$

where the arbitrary constants associated with the two indefinite integrals have been lumped into a single arbitrary constant C.

The solution is not yet in the explicit form $y = F(x)$ for some function F. If we write $C = \ln A$ $(A > 0)$, then the right-hand side becomes $\ln|4 + x^3| + \ln A = \ln(A|4 + x^3|)$, and we can take exponentials of both sides to give the explicit form of the solution as

$$y = A(4 + x^3).$$

If necessary, A can be negative to take into account changes in sign for y or $(4 + x^3)$.

(b) The equation is of the form

$$\frac{dy}{dx} = g(x)\,h(y),$$

where

$$g(x) = -x \quad \text{and} \quad h(y) = 1/(3y).$$

Notice that since $y > 0$, $h(y)$ is never zero.

On dividing through by $h(y) = 1/(3y)$ (i.e. multiplying through by $3y$) and integrating with respect to x, the differential equation becomes

$$\int 3y\,dy = \int -x\,dx.$$

Evaluating the integrals gives

$$\tfrac{3}{2}y^2 = -\tfrac{1}{2}x^2 + C,$$

where C is an arbitrary constant. This is an implicit form of the general solution.

On solving for y (and noting the condition $y > 0$ given above, which determines the sign of the square root), we obtain the explicit general solution

$$y = \sqrt{\tfrac{1}{3}(2C - x^2)}.$$

This can be simplified slightly by writing A in place of $2C$, where A is another arbitrary constant. However, we need to recognise that the formula for y represents a real quantity greater than zero only when the argument of the square root is positive, so we must have $A - x^2 > 0$. This in turn means that A cannot be completely arbitrary, since it must at least be positive. The general solution in this case is therefore

Since $x^2 \geq 0$ for all x, $A - x^2 > 0$ implies $A > x^2 \geq 0$, so A must be positive.

$$y = \sqrt{\tfrac{1}{3}(A - x^2)} \quad (-\sqrt{A} < x < \sqrt{A}),$$

where A is a positive but otherwise arbitrary constant.

The initial condition is $y(0) = 3$, so we substitute $x = 0$ and $y = 3$ into the general solution above. This gives $3 = \sqrt{A/3}$, so $A = 27$, and the required particular solution is

$$y = \sqrt{\tfrac{1}{3}(27 - x^2)} \quad (-3\sqrt{3} < x < 3\sqrt{3}).$$

The method is summarised below.

Procedure 2 Separation of variables

This method applies to separable differential equations, which are of the form

$$\frac{dy}{dx} = g(x)\,h(y).$$

1. Divide both sides by $h(y)$ (where $h(y) \neq 0$), and integrate both sides with respect to x, to obtain
$$\int \frac{1}{h(y)}\,dy = \int g(x)\,dx.$$
2. If possible, perform the integrations to obtain an implicit form of the general solution.
3. If possible, rearrange the formula found in Step 2 to give y in terms of x. This is the explicit (general) solution.

It is a good idea to check, by substitution into the original differential equation, that the function obtained is indeed a solution.

The separation of variables method is useful, but there are some difficulties with it. First, it may not be possible to perform the necessary integrations. Second, the general solution obtained is restricted to those values of y such that $h(y) \neq 0$. Third, it may not be possible to perform the necessary manipulations to obtain an explicit solution.

Of these difficulties, the first can be overcome by use of a numerical method, such as Euler's method (see Subsection 1.5). The second will be discussed shortly. The third will usually also need numerical techniques.

Exercise 7

A mass $m(t)$ of a uranium isotope, which is present in an object at time t, declines over time due to radioactive decay. Its behaviour is modelled by the differential equation

$$\frac{dm}{dt} = -\lambda m \quad (m > 0),$$

where the *decay constant* λ is a positive constant characteristic of the uranium isotope.

This model can be applied to other radioactive substances by selecting the appropriate value of the parameter λ.

Note the condition $m > 0$. You can see that $m = 0$ also satisfies the differential equation.

(a) Find the general solution of this differential equation.

(b) Find the particular solution for which the initial amount of uranium present (at time $t = 0$) is m_0.

The condition $m > 0$ in Exercise 7 arose from the modelling context. This condition enabled us to find the general solution without needing to worry about dividing by zero at Step 1 of the separation of variables method (and hence without needing to restrict the image set further).

We should also note that:

- the separation of variables method requires that $h(y) \neq 0$ and gives a family of solutions containing an arbitrary constant
- the case when $h(y) = 0$ is exceptional and can give extra solutions that may or may not have the same form as the family of general solutions.

The following exercises provide you with some practice at applying the separation of variables method and at completing the general solution for values of y such that $h(y) = 0$.

Exercise 8

Find the general solution of each of the following differential equations.

(a) $\dfrac{dy}{dx} = \dfrac{y-1}{x}$ $(x > 0)$ (b) $\dfrac{dy}{dx} = \dfrac{2y}{x^2+1}$

Exercise 9

The differential equation here is the logistic equation (5), which, as was pointed out earlier, may be used as a model for the size $P(t)$ of a population at time t.

(a) Solve the initial-value problem

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right), \quad P(0) = P_0 \text{ (where } P_0 > 0),$$

where k and M are positive constants.

(*Hint*: For the integral involving P, the solution to Exercise 6(b) should be useful.)

(b) Describe what happens to the solution $P(t)$ as t becomes large.

1.4 Integrating factor method

This subsection presents one final method of analytic solution for *linear* first-order differential equations.

A first-order differential equation for $y = y(x)$ is **linear** if it can be expressed in the form

$$\frac{dy}{dx} + g(x)\,y = h(x), \tag{9}$$

where $g(x)$ and $h(x)$ are given functions.

A linear first-order differential equation is said to be **homogeneous** if $h(x) = 0$ for all x, and **inhomogeneous** or **non-homogeneous** otherwise.

For example, the differential equation

$$\frac{dy}{dx} - x^2 y = x^3$$

is linear, with $g(x) = -x^2$ and $h(x) = x^3$, whereas the equation

$$\frac{dy}{dx} = xy^2$$

is not, due to the presence of the non-linear term y^2.

Exercise 10

Decide whether or not each of the following first-order differential equations is linear.

(a) $\dfrac{dy}{dx} = x \sin x$ (b) $\dot{y} + y^2 = t$

(c) $x\dfrac{dy}{dx} + y = y^2$ (d) $(1 + x^2)\dfrac{dy}{dx} + 2xy = 3x^2$

An important existence and uniqueness theorem for *linear* first-order differential equation guarantees that an initial-value problem has a unique solution.

Theorem 1

If the functions $g(x)$ and $h(x)$ are continuous throughout an interval (a, b) and x_0 belongs to this interval, then the initial-value problem

$$\frac{dy}{dx} + g(x)\, y = h(x), \quad y(x_0) = y_0,$$

has a unique solution throughout the interval.

This includes the possibility that either $a = -\infty$ or $b = \infty$, so the interval might be all of the real line.

This is a very powerful result, since it means that once you have found a solution in a particular interval, that solution will be the *only* one.

The *integrating factor method* is a technique for solving linear differential equations. It derives from the rule for integration by parts or, equivalently, from the product rule for derivatives. To introduce the topic, consider the differential equation

$$(1 + x^2)\frac{dy}{dx} + 2xy = 3x^2. \tag{10}$$

As you saw in Exercise 10(d), this differential equation is linear; but it is not soluble by direct integration or by separation of variables.

Note first that $2x$ (the coefficient of y) is the derivative of $1 + x^2$ (the coefficient of dy/dx). It follows from the product rule that

$$\frac{d}{dx}\big((1 + x^2)y\big) = (1 + x^2)\frac{dy}{dx} + 2xy.$$

The right-hand side of this equation is the same as the left-hand side of equation (10), so we can rewrite the latter as

$$\frac{d}{dx}\big((1 + x^2)y\big) = 3x^2. \tag{11}$$

Now the left-hand side here is just the derivative of $(1 + x^2)y$, so we can apply direct integration to equation (11) to obtain

$$(1 + x^2)y = \int 3x^2\,dx = x^3 + C,$$

where C is an arbitrary constant. Division by $1 + x^2$ then gives the general solution of equation (10) explicitly as

$$y = \frac{x^3 + C}{1 + x^2}.$$

This solution was arrived at by noting that the left-hand side of equation (10) is of the form

$$p\,\frac{dy}{dx} + \frac{dp}{dx}\,y, \tag{12}$$

where $p = 1 + x^2$, and that this form can be re-expressed, using the product rule, as

$$\frac{d}{dx}(py).$$

Linear differential equations need not come in this convenient form. The left-hand side of the equation

$$\frac{dy}{dx} + g(x)\,y = h(x) \tag{13}$$

may not be of the form (12). An *integrating factor* $p = p(x)$ that enables us to transform the left-hand side of equation (13) into the form (12) can be found by writing down the two properties that such a function must satisfy, as follows.

- Multiplying equation (13) by p gives, on the left-hand side,

  $$p\,\frac{dy}{dx} + p\,g(x)\,y.$$

- The left-hand side must be of the form

  $$p\,\frac{dy}{dx} + \frac{dp}{dx}\,y.$$

Comparison of these two expressions shows that p must itself be a particular solution of the differential equation

$$\frac{dp}{dx} = g(x)\,p.$$

This is a homogeneous linear first-order differential equation, and we can solve it by separation of variables. Indeed, following Procedure 2, the equation becomes (for $p \neq 0$)

$$\int \frac{dp}{p} = \int g(x)\,dx.$$

Because any constant multiple of an integrating factor is still an integrating factor, we may assume $p > 0$.

Performing the left-hand integral gives

$$\ln p = \int g(x)\,dx,$$

so

$$p = \exp\left(\int g(x)\,dx\right),$$

which defines the **integrating factor** for equation (13).

When equation (13) is multiplied through by the integrating factor, the resulting differential equation is

$$p(x)\frac{dy}{dx} + p(x)\,g(x)\,y = p(x)\,h(x), \tag{14}$$

the left-hand side of which, by the definition of p, is of the form (12). So equation (14) can be re-expressed, using the product rule, as

$$\frac{d}{dx}(p(x)\,y) = p(x)\,h(x). \tag{15}$$

Direct integration can then be used on equation (15) to try to find the general solution.

This integrating factor method is summarised below.

The definition of p ensures that the left-hand side of equation (14) is of the form (12) since

$$\begin{aligned}\frac{dp}{dx} &= \frac{d}{dx}\left(\exp\left(\int g(x)\,dx\right)\right)\\ &= \exp\left(\int g(x)\,dx\right)g(x)\\ &= p(x)\,g(x).\end{aligned}$$

Procedure 3 Integrating factor method

This method applies to differential equations of the form

$$\frac{dy}{dx} + g(x)\,y = h(x). \tag{16}$$

1. Determine the integrating factor
$$p(x) = \exp\left(\int g(x)\,dx\right). \tag{17}$$
2. Multiply equation (16) by $p(x)$ to recast the differential equation as
$$p(x)\frac{dy}{dx} + p(x)\,g(x)\,y = p(x)\,h(x).$$
3. Rewrite the differential equation as
$$\frac{d}{dx}(p(x)\,y) = p(x)\,h(x).$$
4. Integrate this last equation to obtain
$$p(x)\,y = \int p(x)\,h(x)\,dx.$$
5. Divide through by $p(x)$ to obtain the general solution in explicit form.

The constant of integration is not needed here.

You can, if you wish, check that you have found p correctly by checking that

$$\begin{aligned}p(x)\frac{dy}{dx} + p(x)\,g(x)\,y\\ = \frac{d}{dx}(p(x)\,y),\end{aligned}$$

i.e. by checking that $dp/dx = p(x)\,g(x)$.

The integral in Step 4 will involve an arbitrary constant C.

It is a good idea to check, by substitution into the original equation, that the function obtained is indeed a solution.

As with the separation of variables method, it may not be possible to perform the necessary final integration.

Exercise 11

Find the general solution of each of the following differential equations.

(a) $\dfrac{dy}{dx} - y = e^x \sin x$ (b) $\dfrac{dy}{dx} = \dfrac{y-1}{x}$ $(x > 0)$

Exercise 12

Which method would you use to try to solve each of the following linear first-order differential equations?

(a) $\dfrac{dy}{dx} + x^3 y = x^5$ (b) $\dfrac{dy}{dx} = x \sin x$

(c) $\dfrac{dv}{du} + 5v = 0$ (d) $(1 + x^2)\dfrac{dy}{dx} + 2xy = 1 + x^2$

1.5 Direction fields and Euler's method

The final method of this section is a numerical method that can be used when the analytic methods fail. First, you will see that qualitative information about the solutions of a first-order differential equation may be gleaned directly from the equation itself, without undertaking any form of integration process. The main concept here is the *direction field*, sketches of which usually give a good idea of how the graphs of solutions behave. Direction fields can also be regarded as the starting point for a numerical (i.e. calculational rather than algebraic) method of solution called *Euler's method*.

Consider what can be deduced about solutions of the differential equation

$$\frac{dy}{dx} = f(x, y)$$

from direct observation of this equation.

In Section 1 we encountered the logistic equation

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right), \tag{18}$$

Here we have

$$f(t, P) = kP\left(1 - \frac{P}{M}\right).$$

where k and M are positive constants. In certain circumstances this is a useful mathematical model of population changes, in which $P(t)$ denotes the size of the population at time t. The right-hand side of this equation is equal to zero if either $P = 0$ or $P = M$. Hence, since $dP/dt = 0$ in both cases, each of the constant functions $P = 0$ and $P = M$ is a particular solution of equation (18). Within the model, these solutions correspond to a complete absence of the population ($P = 0$), and an equilibrium population level ($P = M$) for which the proportionate birth and death rates are equal.

Such spotting of constant functions that are particular solutions is useful on occasion but of limited applicability. In general, more useful information can be deduced from the observation that, for any given point (x, y) in the plane, the equation

$$\frac{dy}{dx} = f(x, y) \tag{19}$$

describes the *direction* in which the graph of the particular solution through that point is heading (see Figure 3). This is because if $y = y(x)$ is any solution of the differential equation, then dy/dx is the *gradient* or *slope* of the graph of that function. Equation (19) therefore tells us that $f(x, y)$ represents the slope at (x, y) of the graph of the particular solution that passes through (x, y). If the slope $f(x, y)$ at this point is positive, then the corresponding solution graph is increasing (rising) from left to right through the point (x, y); if the slope is negative, then the graph is decreasing (falling); and if the slope is zero, then the graph is horizontal at the point.

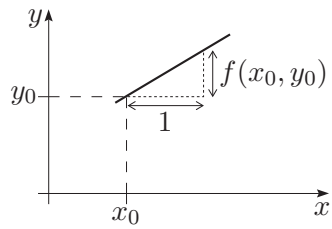


Figure 3 A graphical representation of the slope at the point (x_0, y_0)

For example, if $f(x, y) = x + y$, then the slope at the point $(1, 2)$ is $f(1, 2) = 1 + 2 = 3$, the slope at the point $(2, -7)$ is $f(2, -7) = 2 - 7 = -5$, and the slope at the point $(3, -3)$ is $f(3, -3) = 3 - 3 = 0$.

When looking at $f(x, y)$ in this light, it is referred to as a *direction field*, since it describes a *direction* (slope) for each point (x, y) where $f(x, y)$ is defined.

A **direction field** associates a unique direction to each point within a specified region of the (x, y)-plane. The direction corresponding to the point (x, y) may be thought of as the slope of a short line segment through the point.

In particular, the direction field for the differential equation

$$\frac{dy}{dx} = f(x, y)$$

associates the direction $f(x, y)$ with the point (x, y).

Direction fields can be visualised by constructing the short line segments referred to above for a finite set of points in an appropriate region of the plane, where typically the points are chosen to form a rectangular grid.

Example 4

(a) Part of the direction field for the logistic equation

$$\frac{dP}{dt} = P\left(1 - \frac{P}{1000}\right)$$

This is equation (18) with $k = 1$ and $M = 1000$.

is sketched in Figure 4. Using this diagram, sketch the solution curves that pass through the points

$$(0, 1500),\ (0, 1000),\ (0, 100),\ (0, 0),\ (0, -100).$$

We do not normally need to consider $P < 0$ since populations must be non-negative.

From your results, describe the graphs of particular solutions of the differential equation.

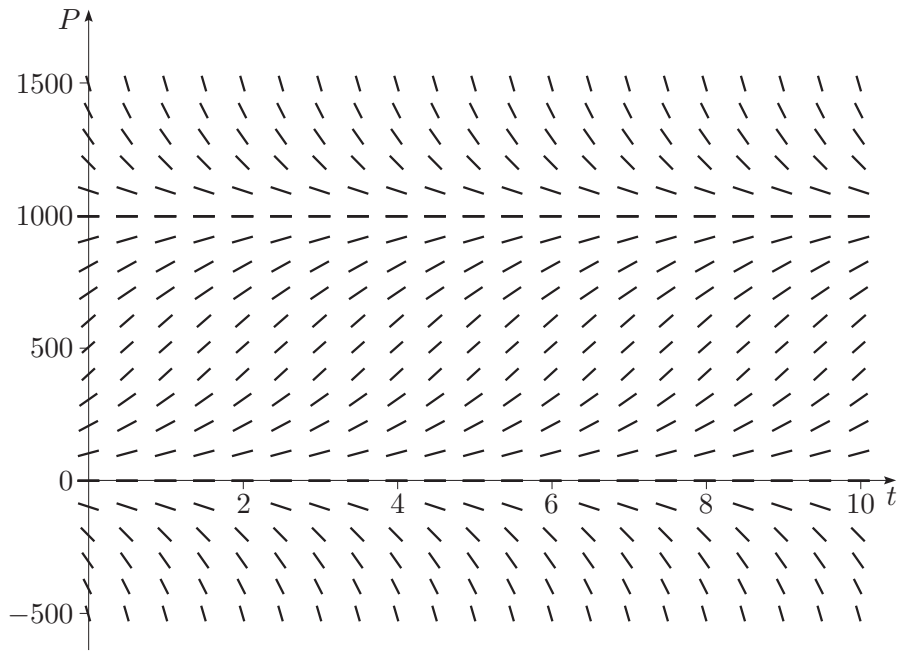


Figure 4 A direction field for the logistic equation

(b) What does your answer to part (a) tell you about the predicted behaviour of a population whose size $P(t)$ at time t is modelled by this logistic equation?

Solution

(a) The slope is shown to be zero at all points on the horizontal lines $P = 0$ and $P = 1000$, so these correspond to constant solutions of the differential equation. (As pointed out earlier in the text, these two solutions can also be spotted directly from the form of the differential equation.)

The graphs of solutions through starting points above the line $P = 1000$ appear to decrease, but at a slower and slower rate, tending from above towards the limit $P = 1000$ as t increases.

The graphs of solutions through starting points in the region $0 < P < 1000$ are increasing, with slope growing before the level $P = 500$ is reached and declining thereafter. For large values of t, these graphs tend from below towards the limit $P = 1000$.

For starting points in the region $P < 0$, the graphs decrease without limit and with steeper and steeper slope.

These various cases are illustrated by typical graphs in Figure 5.

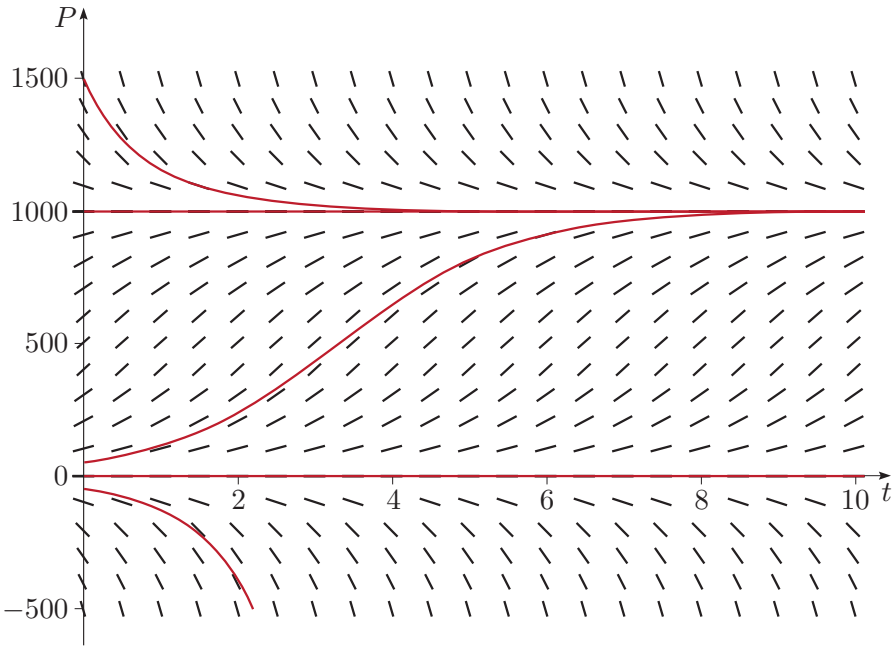


Figure 5 Some solution curves for the logistic equation

(b) If the differential equation is considered as a model of population behaviour, then the region $P < 0$ must be excluded. The analysis above leads to the following predictions for the population.

- If the population is zero at the start, then it remains zero.
- If the population size starts at 1000, then it remains fixed at this level.
- If the population starts at a level higher than 1000, then it declines (more and more gradually) towards 1000.
- If the population starts at a level below 1000 (but above 0), then it increases and eventually tends gradually towards 1000.

Drawing *by hand* precise grids of line segments to represent direction fields is not a good use of your time. However, it is a task that your computer can be programmed to perform. Furthermore, the concept of direction fields helps in constructing approximate numerical solutions for first-order differential equations.

The graphs of particular solutions of a differential equation

$$\frac{dy}{dx} = f(x, y)$$

can be 'mentally sketched' on a diagram of the direction field given by $f(x, y)$. The tangent to the solution curve is always 'parallel to the local slope' of the direction field. While this gives a good visual image of the connection between the direction field and the graph of a solution, it is somewhat short on precision. We could not expect, by this approach, to predict with any accuracy the actual solution to an initial-value problem.

Building on this idea, a solution to an initial-value problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0, \tag{20}$$

may be estimated by calculational means. The direction field diagram helps to explain how this numerical method arises.

Suppose that instead of trying to sketch a solution curve to fit the direction field, we move in a sequence of straight-line steps whose directions are governed by the direction field. The aim is to produce a sequence of points that provide approximate values of the solution function $y(x)$ for the initial-value problem at a sequence of x-values. The steps are constructed as follows.

Corresponding to the given initial condition $y(x_0) = y_0$, there is a point P_0 in the (x, y)-plane with coordinates (x_0, y_0), and this is our starting point. At P_0, the direction field $f(x, y)$ defines a particular slope, namely $f(x_0, y_0)$. We move off from P_0 along a straight line that has this slope, and continue until we have travelled a horizontal distance h to the right of P_0. The point that has now been reached is labelled P_1, as in Figure 6.

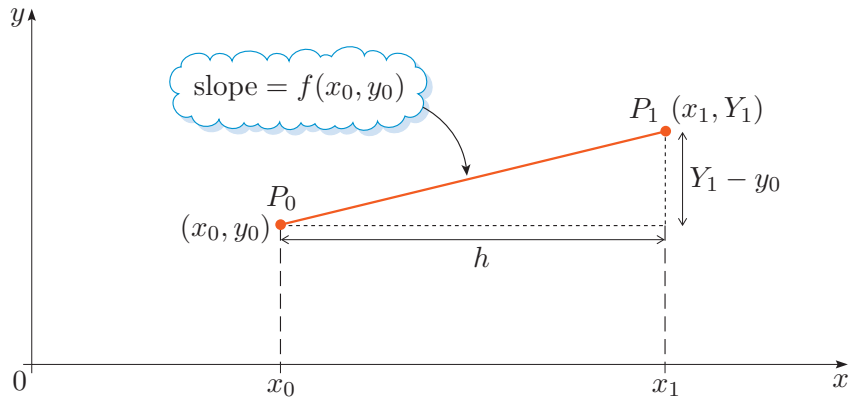


Figure 6 Using the slope at x_0 to estimate the solution at x_1

The reason for using Y_1 here, rather than y_1, will be explained shortly.

The idea is that the point P_1, whose coordinates have been denoted by (x_1, Y_1), provides an approximate value Y_1 of the solution function $y(x)$ at $x = x_1$. Now, unless the solution function follows a straight line as x moves from x_0 to x_1, Y_1 is unlikely to give the exact value of $y(x_1)$. However, the hope is that because we headed off from x_0 along the correct slope, as given by the direction field, Y_1 will be reasonably close to the exact value. Before worrying about accuracy, let us continue with the construction of the points in our sequence.

The next thing that we need to do, before proceeding to the second step in the construction process, is determine formulas for x_1 and Y_1 in terms of x_0, y_0, h and $f(x_0, y_0)$. By the construction described, as the point P_1 is reached from P_0 by taking a step to the right of horizontal length h, we have

$$x_1 = x_0 + h. \tag{21}$$

We can express Y_1 in terms of other quantities by equating two expressions for the slope of the line segment P_0P_1,

$$\frac{Y_1 - y_0}{h} = f(x_0, y_0),$$

and then rearranging to give

$$Y_1 = y_0 + h\, f(x_0, y_0). \tag{22}$$

This completes the first step, and we now take a second step to the right.

The direction of the second step is along the line with slope defined by the direction field at the point P_1, namely $f(x_1, Y_1)$. The second step moves us from P_1 through a further horizontal distance h to the right, to the point labelled P_2, as illustrated in Figure 7. This point provides an approximate value Y_2 of the solution function $y(x)$ at $x = x_2$.

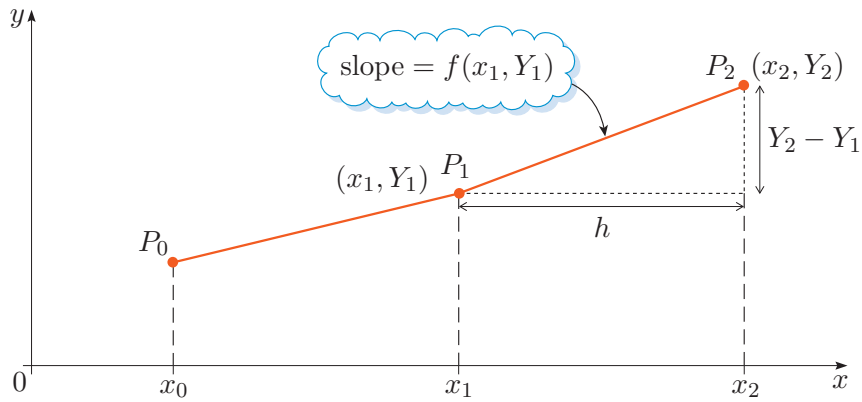


Figure 7 Estimating the value of the solution at x_2

As in the first step, we need now to express the coordinates (x_2, Y_2) of P_2 in terms of x_1, Y_1, h and $f(x_1, Y_1)$. We have

$$x_2 = x_1 + h \tag{23}$$

and (equating two expressions for the slope of the line segment P_1P_2)

$$\frac{Y_2 - Y_1}{h} = f(x_1, Y_1),$$

which can be rearranged to give

$$Y_2 = Y_1 + h\, f(x_1, Y_1). \tag{24}$$

Having carried out two steps of the process, it is possible to see that the same procedure can be applied to construct any number of further steps, and we next generalise to a description of what happens at the $(i+1)$th step, where i represents any non-negative integer.

Suppose that after i steps we have reached the point P_i, with coordinates (x_i, Y_i). For the $(i+1)$th step, we move away from P_i along the line with slope $f(x_i, Y_i)$, as defined by the direction field at P_i. After moving through a horizontal distance h to the right, we reach the point P_{i+1}, whose coordinates are denoted by (x_{i+1}, Y_{i+1}), as illustrated in Figure 8.

The point P_{i+1} provides an approximate value Y_{i+1} of the solution function $y(x)$ at $x = x_{i+1}$.

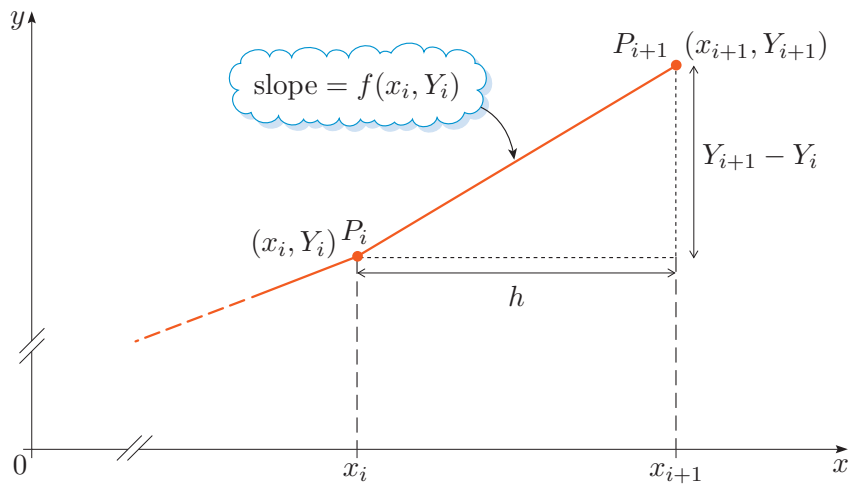


Figure 8 The $(i + 1)$th step of the sequence

Arguing as before, we have

$$x_{i+1} = x_i + h \tag{25}$$

and (equating two expressions for the slope of the line segment P_iP_{i+1})

$$\frac{Y_{i+1} - Y_i}{h} = f(x_i, Y_i),$$

which can be rearranged to give

$$Y_{i+1} = Y_i + h\, f(x_i, Y_i). \tag{26}$$

Note that equations (21) and (23) are the special cases of equation (25) for $i = 0$ and $i = 1$, respectively, and that equation (24) is the special case of equation (26) for $i = 1$. If we also define Y_0 to be equal to the initial value y_0, then equation (22) is the special case of equation (26) for $i = 0$.

To sum up, for the initial-value problem (20), we have a procedure for constructing a sequence of points

$$P_i \text{ with coordinates } (x_i, Y_i) \quad (i = 1, 2, \ldots),$$

where the values of x_i and Y_i for each value of i are determined by the respective formulas (25) and (26). The starting point for the sequence is the point P_0 with coordinates (x_0, Y_0), where $Y_0 = y_0$. Because the procedure is based on the direction field, each Y_i provides an approximation at $x = x_i$ to the value of the solution function $y(x)$ for the initial-value problem. The horizontal distance h by which we move to the right at each stage of the procedure is called the **step size** (or **step length**).

Figure 9 shows the constructed sequence of points, and for comparison includes a curve representing the graph of the exact solution of the initial-value problem (20). This makes clear that the successive points $P_1, P_2, P_3, \ldots$ are only *approximations* to points on the solution curve.

In fact, the situation shown in Figure 9 is typical of the behaviour of the constructed approximations, in that they gradually move further and further from the exact solution curve. This is because at each step, the direction of movement is along the slope of the direction field at $P_i = (x_i, Y_i)$ and not along the slope of the direction field at (x_i, y_i), where $y_i = y(x_i)$ denotes the value of the exact solution function at $x = x_i$; that is, for each x_i, the slope for the next step is defined by a point close to the solution curve rather than by the point exactly on that curve.

The common use of $y_i = y(x_i)$ to represent the exact solution at $x = x_i$ explains why we use a different notation, namely Y_i, for the numerical approximation to $y(x_i)$.

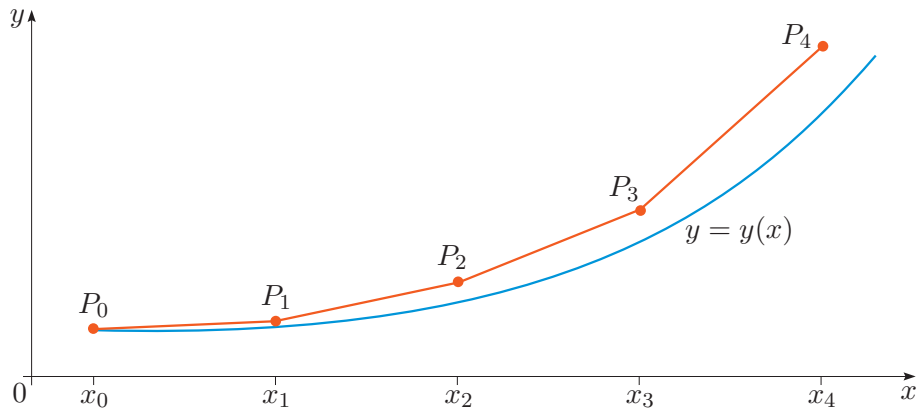


Figure 9 The exact solution and the approximate solution

Nevertheless, the formulas (25) and (26) provide a method for finding approximate solutions to the initial-value problem (20), in terms of numerical estimates $Y_1, Y_2, Y_3, \ldots$ at the respective domain values $x_1, x_2, x_3, \ldots$. This is called *Euler's method*, after Leonhard Euler (Figure 10). It is summarised below.

The accuracy of such approximate solutions, and ways of improving accuracy, will be considered shortly.

Procedure 4 Euler's method

To apply Euler's method to the initial-value problem

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0,$$

proceed as follows.

1. Take x_0 and $Y_0 = y_0$ as starting values, choose a step size h, and set $i = 0$.
2. Calculate the x-coordinate x_{i+1}, using the formula
$$x_{i+1} = x_i + h.$$
3. Calculate a corresponding approximation Y_{i+1} to $y(x_{i+1})$, using the formula
$$Y_{i+1} = Y_i + h\, f(x_i, Y_i).$$
4. If further approximate values are required, increase i by 1 and return to Step 2.

Figure 10 Leonhard Euler (1707–1783) was one of the most prolific mathematicians of all time. (His surname is pronounced 'oiler'.) He first devised this method in order to compute the orbit of the Moon.

Example 5

Consider the initial-value problem

$$\frac{dy}{dx} = x + y, \quad y(0) = 1.$$

Use Euler's method, with step size $h = 0.2$, to obtain an approximation to $y(1)$.

Solution

We have $x_0 = 0$, $Y_0 = y_0 = 1$, and $f(x_i, Y_i) = x_i + Y_i$. The step size is given as $h = 0.2$. Equation (25) with $i = 0$ gives

$$x_1 = x_0 + h = 0 + 0.2 = 0.2,$$

and equation (26) with $i = 0$ gives

$$Y_1 = Y_0 + h\, f(x_0, Y_0) = 1 + 0.2 \times (0 + 1) = 1.2.$$

For the second step, we have (from equation (25) with $i = 1$)

$$x_2 = x_1 + h = 0.2 + 0.2 = 0.4,$$

and (from equation (26) with $i = 1$)

$$Y_2 = Y_1 + h\, f(x_1, Y_1) = 1.2 + 0.2 \times (0.2 + 1.2) = 1.48.$$

If more than a couple of steps of such a calculation have to be computed by hand, then it is a good idea to lay out the calculation as a table. In this case, by continuing as above and putting i in turn equal to 2, 3 and 4, we obtain Table 1.

Table 1

i	x_i	Y_i	$f(x_i, Y_i) = x_i + Y_i$	$Y_{i+1} = Y_i + h\, f(x_i, Y_i)$
0	0	1	1	1.2
1	0.2	1.2	1.4	1.48
2	0.4	1.48	1.88	1.856
3	0.6	1.856	2.456	2.347 2
4	0.8	2.347 2	3.147 2	2.976 64
5	1.0	2.976 64		

(After each value of Y_{i+1} has been calculated from the formula and entered in the last column, it is transferred to the Y_i column in the next row. Once the value at $x = 1.0$ has been found, no further calculations are necessary.)

So at $x = 1$, Euler's method with step size $h = 0.2$ gives the approximation $y(1) \simeq 2.976\,64$.

Exercise 13

Use Euler's method, with step size $h = 0.2$, to obtain an approximation to $y(1)$ for the initial-value problem

$$\frac{dy}{dx} = y, \quad y(0) = 1.$$

The solution to the initial-value problem in Example 5 is in fact known exactly, and is $y = 2e^x - x - 1$. Putting $x = 1$, this gives

$$y(1) = 2e - 1 - 1 = 3.436\,56,$$

correct to five decimal places. This value may be compared with the approximation 2.976 64 for $y(1)$ obtained by Euler's method in Example 5, and the comparison indicates that the approximation is not at all accurate. Indeed, the *absolute error* in this case is

$$|2.976\,64 - 3.436\,56| = 0.459\,92,$$

which is about 13% of the exact value, and indeed not even one decimal place accuracy is achieved.

The absolute error is defined to be the magnitude of the difference between the approximate value and the exact value.

Similarly, the other values Y_i $(i = 1, 2, 3, 4)$ found in Example 5 contain significant absolute errors when considered as approximations to the corresponding exact values $y_i = y(x_i)$. This is illustrated in general terms in Figure 9, where the absolute error in approximation Y_i is the vertical distance from the point P_i to the point directly below it on the exact solution curve. As shown here, and for reasons given earlier, the absolute error tends to increase as more and more steps are taken.

The realisation that Euler's method can produce values that are poor approximations to the exact solution to an initial-value problem invites us to ask whether the accuracy of the approximations can be improved using this method. In fact, it is not hard to see that improvements in accuracy ought to be achieved by *reducing the step size h*. Our earlier prescription for constructing the sequence of points $P_1, P_2, P_3, \ldots$ from the starting point P_0 and the given direction field amounts loosely to saying 'match the direction of the solution curve at the current point, take a step, then adjust direction so as to try not to move further away from the curve'. It seems natural, therefore, that the approximations will improve if we reduce the size of the steps taken and correspondingly 'adjust direction' more frequently. This is illustrated for a hypothetical case in Figure 11.

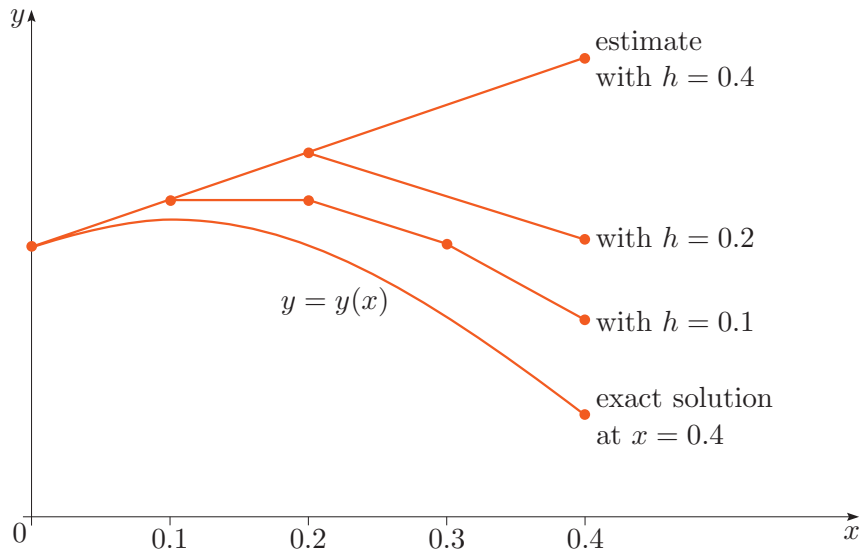


Figure 11 Comparing the approximate solution with the exact solution as h decreases

In fact, it can be shown that the accuracy of Euler's method does indeed usually improve when we take a smaller step size.

To demonstrate this, consider the initial-value problem from Exercise 13. This has the exact solution $y = e^x$ (as you can verify), and its value at $x = 1$ is $y(1) = e = 2.718\,282$, to six decimal places. In Exercise 13 you showed that with a step size $h = 0.2$, Euler's method gives the approximation 2.488 32 for $y(1)$. Table 2 shows the corresponding results (to six decimal places) obtained when we apply Euler's method to this same initial-value problem but with successively smaller step sizes h.

In Exercise 13, where $h = 0.2$, the value of $y(1)$ was approximated by Y_5. From the column for 'Number of steps' in Table 2, you can see that $y(1)$ is approximated by:

Y_{10} when $h = 0.1$;
Y_{100} when $h = 0.01$;
Y_{1000} when $h = 0.001$;
Y_{10000} when $h = 0.0001$.

Table 2

h	Approximation to $y(1)$	Absolute error	Number of steps
0.1	2.593 742	0.124 539	10
0.01	2.704 814	0.013 468	100
0.001	2.716 924	0.001 358	1000
0.0001	2.718 146	0.000 136	10000

As expected, the absolute errors in the third column of the table become progressively smaller as h is reduced.

Looking more carefully at these absolute errors, we notice that they seem to tend towards a sequence in which each number is a tenth of the previous one. Since each value of h in the table is a tenth of the previous one, this suggests that

absolute error is approximately proportional to step size h.

This turns out to be a general property of Euler's method, for sufficiently small values of the step size. So not only do we know that accuracy can be

improved by decreasing the step size h, but this general property also tells us that by making h small enough, the absolute error in an approximation can be made as small as desired. In other words, the absolute error approaches the limit zero as h approaches zero (as you might have expected from the intuitive argument preceding Figure 11).

Exercise 14

Suppose that when Euler's method is applied to the problem in Exercise 13, the absolute error in approximating $y(1)$ is proportional to the step size h, for sufficiently small h.

Use the last row of Table 2 to estimate the constant of proportionality, k say, and hence estimate the step size required to compute $y(1)$ correct to five decimal places (i.e. so that the absolute error is less than 5×10^{-6}).

A few words of caution are necessary at this point. Although the absolute error can be made as small as we please by making the step size h sufficiently small, this is valid *only* if the arithmetic is performed using sufficient decimal places. Where a calculator or computer is involved, the number of decimal places that can be used is limited, and as a result *rounding errors* may be introduced into the calculations. After a certain point, any increase in accuracy brought about by reducing the size of h may be swamped by these rounding errors.

Moreover, rounding errors are not the only problem. Before concluding that h should always be chosen to be very small, we must also consider the cost of this additional accuracy. Now, by *cost* is meant the effort involved, which can be measured in a variety of ways; commonly for iterative methods (such as Euler's method) it is measured by the number of steps taken. In general for numerical methods, the greater the accuracy required, the greater the cost. To illustrate this, look back at Table 2. The last column of the table shows how the number of steps required for the calculation goes up in inverse proportion to the step size: to move from $x = 0$ to $x = 1$, it takes 10 steps with step size $h = 0.1$, 100 steps with step size $h = 0.01$, and so on. Since, for sufficiently small h, the error in Euler's method is approximately proportional to the step size, it follows that for this method a tenfold improvement in accuracy is paid for by a tenfold increase in the number of steps required.

In general, to move from a to b (where $b > a$) with step size h takes $(b - a)/h$ steps.

So for Euler's method and similar methods, the choice of step size has to be based on a compromise between the two opposing requirements of accuracy and cost. There are better numerical methods for solving initial-value problems that are considerably more *efficient* than Euler's method. In fact, Euler's method is not suitable for high-accuracy work. Its virtue lies rather in its simplicity and its clear illustration of the basic principles of how differential equations may be solved numerically.

Greater *efficiency* means that the same or better numerical accuracy is achieved with fewer numerical computations.

In any practical case, calculations of the type described in this subsection are ideally suited to being performed on a computer.

Exercise 15

(a) Without plotting the direction field, say what you can about the slopes defined by the differential equation

$$\frac{dy}{dx} = f(x, y) = y + x^2.$$

(b) Verify that your conclusions are consistent with the direction field diagram in Figure 12.

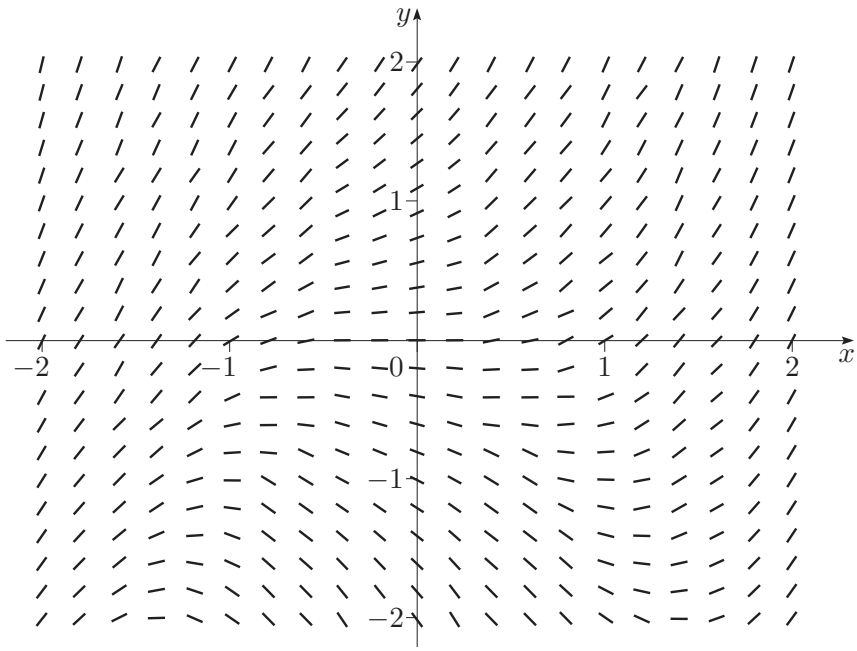


Figure 12 A direction field

(c) On the basis of the direction field, what can be said about the graphs of solutions of the differential equation?

(d) Write down the formulas required in order to apply Euler's method to the initial-value problem

$$\frac{dy}{dx} = y + x^2, \quad y(-1) = -0.2,$$

using a step size $h = 0.1$.

2 Homogeneous second-order differential equations

You will recall that a *particular solution* of a first-order differential equation is obtained by applying a *single* condition (known as an *initial condition*) to the general solution in order to find a particular value of the single arbitrary constant. In the case of a second-order differential equation, a particular solution is obtained by applying *two* conditions to the general solution in order to find particular values of the *two* arbitrary constants. The following example illustrates this.

Example 6

Suppose that a car is travelling with constant acceleration a along a straight road. If, at time t, its distance from a fixed point is s, then its velocity is given by ds/dt, its acceleration is given by d^2s/dt^2, and its motion is modelled by

$$\frac{d^2s}{dt^2} = a. \tag{27}$$

If the car is initially stationary at position $s = 0$ and thereafter has a constant acceleration of $2\,\mathrm{m\,s^{-2}}$, how long does it take for the car to attain a velocity of $30\,\mathrm{m\,s^{-1}}$, and what distance has it travelled in that time?

Solution

Integrating equation (27) leads to

$$\frac{ds}{dt} = at + C \quad \text{and} \quad s = \tfrac{1}{2}at^2 + Ct + D,$$

where C and D are arbitrary constants. To find these constants (and hence answer the questions asked), we need to make use of the conditions given. These are that the car is initially stationary (i.e. $ds/dt = 0$ when $t = 0$) at position $s = 0$ (i.e. $s = 0$ when $t = 0$). The first of these conditions together with the equation $ds/dt = at + C$ tells us that $C = 0$. With $C = 0$, the second equation becomes $s = \frac{1}{2}at^2 + D$, and this together with the second condition tells us that $D = 0$.

Therefore when $a = 2$, we have

$$\frac{d^2s}{dt^2} = 2, \quad \frac{ds}{dt} = 2t, \quad s = t^2.$$

So the velocity is $ds/dt = 30$ when $2t = 30$, that is, after 15 seconds, and in this time the car has travelled a distance $s = 15^2$, that is, 225 metres.

In mathematical modelling, the parameters of the model, in this case a, are retained for as long as possible, in order to generate results that could be used in other situations, with different values.

The solution of second-order differential equations is rarely as easy as the solution of equation (27) above. In fact, the approach of repeated direct integration works for only some equations of the form

$$\frac{d^2y}{dx^2} = f(x).$$

Most second-order differential equations cannot be solved by analytic methods at all, and numerical methods have to be employed instead. However, there is one important class of second-order differential equations that can be solved by analytic means; this is the topic of this section, and we introduce it next.

2.1 Linear constant-coefficient differential equations

The rest of this unit considers *linear constant-coefficient* second-order differential equations. But what exactly do the terms 'linear' and 'constant-coefficient' mean in this context? The answers lie in the following definitions.

Compare the definitions for first-order equations in Subsection 1.4. The important feature is the *linear* combination of y and its derivatives on the left-hand side.

A second-order differential equation for $y = y(x)$ is **linear** if it can be expressed in the form

$$a(x)\frac{d^2y}{dx^2} + b(x)\frac{dy}{dx} + c(x)\,y = f(x), \tag{28}$$

where $a(x)$, $b(x)$, $c(x)$ and $f(x)$ are given continuous functions.

A linear second-order differential equation is **constant-coefficient** if the functions $a(x)$, $b(x)$ and $c(x)$ are all constant, so that the equation is of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x), \tag{29}$$

If $a = 0$, then the equation is first-order.

where $a \neq 0$.

A linear constant-coefficient second-order differential equation is said to be **homogeneous** if $f(x) = 0$ for all x, and **inhomogeneous** (or **non-homogeneous**) otherwise.

Linear constant-coefficient second-order differential equations can be written in other ways. For example, we can divide equation (29) through by a to obtain an equation of the form

$$\frac{d^2y}{dx^2} + \beta\frac{dy}{dx} + \gamma y = \phi(x),$$

and this more closely resembles the definition of linear first-order differential equations from Section 1.

Exercise 16

Consider the following second-order differential equations.

(i) $\dfrac{d^2y}{dx^2} = x^2$ (ii) $3\dfrac{d^2y}{dx^2} + 4\dfrac{dy}{dx} + y = x^2$ (iii) $3\dfrac{d^2y}{dx^2} + 4\dfrac{dy}{dx} + y = 0$

(iv) $xy'' + x^2y = 0$ (v) $2y\dfrac{d^2y}{dx^2} + 4y = 3\dfrac{dy}{dx}$

(vi) $2\dfrac{d^2t}{d\theta^2} + 3\dfrac{dt}{d\theta} + 4t = \sin\theta$ (vii) $\ddot{x} = -4t$ (viii) $\ddot{x} = -4x$

(a) Which of the equations are both linear and constant-coefficient?

(b) Which of the linear constant-coefficient equations are homogeneous?

(c) For each equation, identify the dependent and independent variables.

One of the main reasons for concentrating on linear constant-coefficient differential equations is that there is a large body of theory on which we can call in order to solve them. Subsection 2.2 illustrates this.

Principle of superposition

A key theoretical result will turn out to be extremely useful throughout this module. This is known as the *principle of superposition*, and it is a fundamental property of linear differential equations.

Suppose that we have a solution $y_1(x)$ of

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f_1(x),$$

Here a, b and c can be functions of x.

and a solution $y_2(x)$ of

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f_2(x).$$

Then we claim that the linear combination $k_1y_1 + k_2y_2$, where k_1 and k_2 are constants, is a solution of

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = k_1\,f_1(x) + k_2\,f_2(x). \tag{30}$$

In fact, this is easy to see, for if we substitute $k_1y_1 + k_2y_2$ directly into equation (30), we obtain

$$\begin{aligned}
&a\frac{d^2}{dx^2}(k_1y_1 + k_2y_2) + b\frac{d}{dx}(k_1y_1 + k_2y_2) + c(k_1y_1 + k_2y_2)\\
&\quad = a\left(k_1\frac{d^2y_1}{dx^2} + k_2\frac{d^2y_2}{dx^2}\right) + b\left(k_1\frac{dy_1}{dx} + k_2\frac{dy_2}{dx}\right) + c(k_1y_1 + k_2y_2)\\
&\quad = k_1\left(a\frac{d^2y_1}{dx^2} + b\frac{dy_1}{dx} + cy_1\right) + k_2\left(a\frac{d^2y_2}{dx^2} + b\frac{dy_2}{dx} + cy_2\right)\\
&\quad = k_1\,f_1(x) + k_2\,f_2(x),
\end{aligned}$$

as required.

We summarise this important result as a theorem.

Theorem 2 Principle of superposition

If $y_1(x)$ is a solution of the linear second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f_1(x),$$

and $y_2(x)$ is a solution of the linear second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f_2(x)$$

(with the same left-hand side), then the function

$$y(x) = k_1\, y_1(x) + k_2\, y_2(x),$$

where k_1 and k_2 are constants, is a solution of the differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = k_1\, f_1(x) + k_2\, f_2(x).$$

2.2 Method of solution

This subsection develops a method for solving homogeneous linear constant-coefficient second-order differential equations, that is, equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0, \tag{31}$$

where a, b, c are constants and $a \neq 0$.

To see how this method arises, consider the *first*-order differential equation

$$b\frac{dy}{dx} + cy = 0, \tag{32}$$

where b and c are constants and $b \neq 0$. This homogeneous linear equation, which can be solved using the integrating factor method, has a general solution of the form $y = Ae^{\lambda x}$, where A is an arbitrary constant and λ is some fixed constant. To find λ, we can substitute $y = Ae^{\lambda x}$ and $dy/dx = \lambda Ae^{\lambda x}$ into equation (32) to give

$$b\frac{dy}{dx} + cy = b\lambda Ae^{\lambda x} + cAe^{\lambda x} = (b\lambda + c)Ae^{\lambda x}.$$

Therefore, for $y = Ae^{\lambda x}$ to be a solution, $(b\lambda + c)Ae^{\lambda x}$ must be zero for all x. Since A is arbitrary and $e^{\lambda x} > 0$ for all x, we must have $b\lambda + c = 0$, that is, $\lambda = -c/b$.

This useful idea of substituting $y = Ae^{\lambda x}$ as a possible solution can be applied to equation (31) as well. Let us suppose that equation (31) has a solution of the form $y = Ae^{\lambda x}$, for some value of λ. If so, then

$dy/dx = \lambda Ae^{\lambda x}$ and $d^2y/dx^2 = \lambda^2 Ae^{\lambda x}$, and substituting into the left-hand side of equation (31) gives

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = a\lambda^2 Ae^{\lambda x} + b\lambda Ae^{\lambda x} + cAe^{\lambda x}$$
$$= (a\lambda^2 + b\lambda + c)Ae^{\lambda x}.$$

Hence $y = Ae^{\lambda x}$ is indeed a solution of equation (31), for any value of A, provided that λ satisfies

$$a\lambda^2 + b\lambda + c = 0.$$

Note that the discussion here applies irrespective of whether λ is real or complex. The consequences of λ being complex are explained later.

This equation plays such an important role in solving linear constant-coefficient second-order differential equations that it is given a special name.

The **auxiliary equation** of the homogeneous linear constant-coefficient second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0$$

is the quadratic equation

$$a\lambda^2 + b\lambda + c = 0. \qquad (33)$$

The auxiliary equation is sometimes called the *characteristic equation*.

The auxiliary equation is obtained from the differential equation by replacing y by 1, $\frac{dy}{dx}$ by λ, and $\frac{d^2y}{dx^2}$ by λ^2.

Example 7

Write down the auxiliary equation of the differential equation

$$3\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + 4y = 0.$$

Solution

The auxiliary equation is

$$3\lambda^2 - 2\lambda + 4 = 0.$$

Exercise 17

Write down the auxiliary equation of each of the following differential equations.

(a) $\frac{d^2y}{dx^2} - 5\frac{dy}{dx} + 6y = 0$ (b) $y'' - 9y = 0$ (c) $\ddot{x} + 2\dot{x} = 0$

Now, so far, we know that $y = Ae^{\lambda x}$ is a solution of equation (31) provided that λ satisfies its auxiliary equation. But the auxiliary equation is a quadratic equation with real coefficients, so it has two roots (which in general are distinct). These two roots, λ_1 and λ_2 say, give two solutions $y_1 = Ce^{\lambda_1 x}$ and $y_2 = De^{\lambda_2 x}$ of equation (31), where C and D are arbitrary constants.

If $\lambda_1 = \lambda_2$, then we obtain only one solution. This case is dealt with separately below.

Example 8

(a) Write down the auxiliary equation of the differential equation

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0,$$

and find its roots λ_1 and λ_2.

(b) Deduce that $y_1 = Ce^x$ and $y_2 = De^{2x}$ are both solutions of the differential equation, for any values of the two constants C and D.

(c) Show that $y = Ce^x + De^{2x}$ is also a solution of the differential equation, for any values of the two constants C and D.

Solution

(a) The auxiliary equation is

$$\lambda^2 - 3\lambda + 2 = 0.$$

This equation may be solved, for example, by factorising in the form $(\lambda - 1)(\lambda - 2) = 0$, to give the two roots $\lambda_1 = 1$ and $\lambda_2 = 2$.

Using the formula

$$\lambda_1, \lambda_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

produces the same answer. It does not matter which of the roots is called λ_1 and which is called λ_2.

(b) Since $\lambda_1 = 1$ and $\lambda_2 = 2$ are the roots of the auxiliary equation, $y_1 = Ce^x$ and $y_2 = De^{2x}$ are solutions of the differential equation, for any values of C and D.

(c) To show that $y = Ce^x + De^{2x}$ is a solution of the differential equation, we differentiate and substitute into the differential equation. Differentiating to obtain the first and second derivatives of y gives

$$\frac{dy}{dx} = Ce^x + 2De^{2x} \quad \text{and} \quad \frac{d^2y}{dx^2} = Ce^x + 4De^{2x}.$$

Substituting these into the left-hand side of the differential equation gives

$$\begin{aligned}
&\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y \\
&= \left(Ce^x + 4De^{2x}\right) - 3\left(Ce^x + 2De^{2x}\right) + 2\left(Ce^x + De^{2x}\right) \\
&= C(1 - 3 + 2)e^x + D(4 - 6 + 2)e^{2x} \\
&= 0.
\end{aligned}$$

Hence $y = Ce^x + De^{2x}$ is a solution of the differential equation, for any values of C and D.

It can be seen that if λ_1 and λ_2 are *distinct* roots of the auxiliary equation of a homogeneous linear constant-coefficient second-order differential equation, then any solution of the form

$$y = Ce^{\lambda_1 x} + De^{\lambda_2 x}, \tag{34}$$

for some choice of constants C and D, is also a solution. Furthermore, it can be shown that a solution of this form, for distinct roots of the auxiliary equation, and involving two arbitrary constants, is the general solution of the homogeneous linear second-order differential equation.

Exercise 18

Use the auxiliary equation to find the general solution of each of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} + 5\dfrac{dy}{dx} + 6y = 0$ (b) $2\dfrac{d^2y}{dx^2} + 3\dfrac{dy}{dx} = 0$ (c) $\dfrac{d^2z}{du^2} - 4z = 0$

We now consider an example where the two roots of the auxiliary equation are *equal*, in which case the above method does not work! Indeed, in light of the earlier discussion, you might expect the solution always to be of the form $y = Ae^{\lambda_1 x} + Be^{\lambda_2 x}$, where A and B are arbitrary constants. But if $\lambda_1 = \lambda_2$, this reduces to $y = (A + B)e^{\lambda_1 x} = Ce^{\lambda_1 x}$, where $C = A + B$ is a *single* arbitrary constant, so this cannot be the general solution of a *second*-order differential equation, which requires two arbitrary constants.

Example 9

(a) Write down the auxiliary equation of the differential equation

$$\frac{d^2y}{dx^2} + 6\frac{dy}{dx} + 9y = 0,$$

and find its roots λ_1 and λ_2.

(b) Deduce that $y_1 = Ce^{-3x}$ is a solution of the differential equation, for any value of the constant C.

(c) Show that $y_2 = Dxe^{-3x}$ is also a solution, for any value of the constant D.

(d) Deduce that $y = (C + Dx)e^{-3x}$ is also a solution of the differential equation, for any values of the two constants C and D.

Solution

(a) The auxiliary equation is

$$\lambda^2 + 6\lambda + 9 = 0.$$

The left-hand side is the perfect square $(\lambda + 3)^2$, so the auxiliary equation has equal roots $\lambda_1 = \lambda_2 = -3$.

Note that the 'other' root $\lambda_2 = -3$ gives the same solution.

(b) Since $\lambda_1 = -3$ is a root of the auxiliary equation, $y_1 = Ce^{-3x}$ is a solution of the differential equation, for any value of C.

(c) To show that $y_2 = Dxe^{-3x}$ is a solution of the differential equation, we differentiate and substitute into the differential equation.
Differentiating to obtain the first and second derivatives of y_2 gives

Here we are using the product rule for differentiation.

$$\frac{dy_2}{dx} = De^{-3x} + Dx\left(-3e^{-3x}\right) = D(1-3x)e^{-3x},$$
$$\frac{d^2y_2}{dx^2} = -3De^{-3x} + D(1-3x)(-3e^{-3x}) = D(-6+9x)e^{-3x}.$$

Substituting these into the left-hand side of the differential equation gives

$$\begin{aligned}\frac{d^2y_2}{dx^2} + 6\frac{dy_2}{dx} + 9y_2 &= D(-6+9x)e^{-3x} + 6D(1-3x)e^{-3x} + 9Dxe^{-3x}\\ &= D(-6+6)e^{-3x} + D(9-18+9)xe^{-3x}\\ &= 0.\end{aligned}$$

Hence $y_2 = Dxe^{-3x}$ is a solution of the differential equation, for any value of D.

(d) Since $y_1 = Ce^{-3x}$ and $y_2 = Dxe^{-3x}$ are both solutions of the differential equation, the principle of superposition tells us that so is $y = Ce^{-3x} + Dxe^{-3x} = (C+Dx)e^{-3x}$, for any values of C and D.

The solution in Example 9 is of the form $y = Ce^{\lambda_1 x} + Dxe^{\lambda_1 x}$. The extra x in the second term, $Dxe^{\lambda_1 x}$, is needed, in this special case, to incorporate the second arbitrary constant required by the general solution of a second-order differential equation.

In general, when $\lambda_1 = \lambda_2$, $y = xe^{\lambda_1 x}$ is a solution of

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0, \tag{35}$$

where a, b, c are constants and $a \neq 0$.

To see this, differentiate twice to obtain

$$\frac{dy}{dx} = e^{\lambda_1 x} + \lambda_1 xe^{\lambda_1 x} = (1+\lambda_1 x)e^{\lambda_1 x},$$
$$\frac{d^2y}{dx^2} = \lambda_1 e^{\lambda_1 x} + \lambda_1(1+\lambda_1 x)e^{\lambda_1 x} = (2\lambda_1 + \lambda_1^2 x)e^{\lambda_1 x},$$

and substitute into the left-hand side of equation (35) to obtain

$$\begin{aligned}&a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy\\ &= a\left((2\lambda_1 + \lambda_1^2 x)e^{\lambda_1 x}\right) + b\left((1+\lambda_1 x)e^{\lambda_1 x}\right) + c\left(xe^{\lambda_1 x}\right)\\ &= e^{\lambda_1 x}\left(a(2\lambda_1 + \lambda_1^2 x) + b(1+\lambda_1 x) + cx\right)\\ &= e^{\lambda_1 x}\left((2a\lambda_1 + b) + (a\lambda_1^2 + b\lambda_1 + c)x\right).\end{aligned} \tag{36}$$

Since λ_1 is the solution of the auxiliary equation, we have $a\lambda_1^2 + b\lambda_1 + c = 0$. Also, the formula method for solving the auxiliary equation $a\lambda^2 + b\lambda + c = 0$ gives

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Since in this case we have equal roots, we must have $b^2 - 4ac = 0$, so $\lambda_1 = -b/2a$, that is, $2a\lambda_1 + b = 0$. Thus the right-hand side of equation (36) is zero, and $y = xe^{\lambda_1 x}$ is indeed a solution of equation (35). Therefore when $\lambda_1 = \lambda_2$, by the principle of superposition,

$$y = Ce^{\lambda_1 x} + Dxe^{\lambda_1 x} = (C + Dx)e^{\lambda_1 x}, \tag{37}$$

where C and D are arbitrary constants, is always a solution of equation (35).

Exercise 19

Use the auxiliary equation to find the general solution of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + y = 0$ (b) $\ddot{s} - 4\dot{s} + 4s = 0$

Equations (34) and (37) give us the general solution of equation (35) for the cases where the roots λ_1 and λ_2 of the auxiliary equation are distinct and equal, respectively. However, the distinct roots of a quadratic equation may not be real – they could consist of a pair of complex conjugate roots $\lambda_1 = \alpha + \beta i$ and $\lambda_2 = \alpha - \beta i$. If the auxiliary equation has such a pair of roots, we can still write the general solution in the form

Recall that the complex conjugate of the complex number $\alpha + \beta i$ is $\alpha - \beta i$.

$$y = Ae^{\lambda_1 x} + Be^{\lambda_2 x} = Ae^{(\alpha+\beta i)x} + Be^{(\alpha-\beta i)x},$$

but we now have a complex-valued solution.

You will soon see why we use A and B for the arbitrary constants (rather than our usual choice of C and D).

Since equation (35) has real coefficients, we would like a real-valued solution. In order to achieve this, we need to allow A and B to be complex. Then we can use Euler's formula, which tells us that

Euler's formula is given in the Handbook.

$$e^{i\beta x} = \cos\beta x + i\sin\beta x \quad \text{and} \quad e^{-i\beta x} = \cos\beta x - i\sin\beta x.$$

Now

$$\begin{aligned} y &= Ae^{\lambda_1 x} + Be^{\lambda_2 x} \\ &= Ae^{(\alpha+\beta i)x} + Be^{(\alpha-\beta i)x} \\ &= Ae^{\alpha x}e^{i\beta x} + Be^{\alpha x}e^{-i\beta x} \\ &= Ae^{\alpha x}(\cos\beta x + i\sin\beta x) + Be^{\alpha x}(\cos\beta x - i\sin\beta x) \\ &= e^{\alpha x}\big((A+B)\cos\beta x + (Ai - Bi)\sin\beta x\big) \\ &= e^{\alpha x}(C\cos\beta x + D\sin\beta x), \end{aligned}$$

The constants in the final expression are now C and D, in keeping with our previous solutions.

where $C = A + B$ and $D = (A - B)i$. Provided that any initial conditions are real-valued, C and D are real, and this is the required real-valued solution containing two arbitrary constants.

Example 10

(a) Write down the auxiliary equation of the differential equation

$$\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 13y = 0,$$

and show that its roots are $\lambda_1 = 3 + 2i$ and $\lambda_2 = 3 - 2i$.

(b) Confirm that $y_1 = e^{3x}\cos 2x$ and $y_2 = e^{3x}\sin 2x$ are both solutions of the differential equation.

(c) Deduce that $y = e^{3x}(C\cos 2x + D\sin 2x)$ is also a solution of the differential equation, for any values of the two constants C and D.

Solution

(a) The characteristic equation is

$$\lambda^2 - 6\lambda + 13 = 0.$$

The formula method gives

$$\lambda = \frac{6 \pm \sqrt{36 - 4 \times 1 \times 13}}{2} = \frac{6 \pm \sqrt{-16}}{2} = 3 \pm 2i,$$

so the two complex conjugate roots are $\lambda_1 = 3 + 2i$ and $\lambda_2 = 3 - 2i$.

With the previous notation we have $\alpha = 3$ and $\beta = 2$.

(b) To confirm that $y_1 = e^{3x}\cos 2x$ is a solution of the differential equation, we differentiate and substitute into the differential equation. Differentiating to obtain the first and second derivatives of y_1 gives

$$\frac{dy_1}{dx} = 3e^{3x}\cos 2x + e^{3x}(-2\sin 2x) = e^{3x}(3\cos 2x - 2\sin 2x),$$

$$\begin{aligned}\frac{d^2y_1}{dx^2} &= 3e^{3x}(3\cos 2x - 2\sin 2x) + e^{3x}(-6\sin 2x - 4\cos 2x)\\ &= e^{3x}(5\cos 2x - 12\sin 2x).\end{aligned}$$

Substituting these into the left-hand side of the differential equation gives

$$\begin{aligned}\frac{d^2y_1}{dx^2} - 6\frac{dy_1}{dx} + 13y_1 &= e^{3x}(5\cos 2x - 12\sin 2x)\\ &\quad - 6e^{3x}(3\cos 2x - 2\sin 2x) + 13e^{3x}\cos 2x\\ &= e^{3x}[(5 - 18 + 13)\cos 2x + (-12 + 12)\sin 2x]\\ &= 0.\end{aligned}$$

Hence $y_1 = e^{3x}\cos 2x$ is a solution.

Similarly, for $y_2 = e^{3x}\sin 2x$ we have

$$\frac{dy_2}{dx} = e^{3x}(2\cos 2x + 3\sin 2x), \quad \frac{d^2y_2}{dx^2} = e^{3x}(12\cos 2x + 5\sin 2x),$$

and substituting into the left-hand side of the differential equation gives

$$\begin{aligned}\frac{d^2y_2}{dx^2} - 6\frac{dy_2}{dx} + 13y_2 &= e^{3x}[(12 - 12)\cos 2x + (5 - 18 + 13)\sin 2x]\\ &= 0.\end{aligned}$$

Hence $y_2 = e^{3x}\sin 2x$ is also a solution.

(c) Since $y_1 = e^{3x}\cos 2x$ and $y_2 = e^{3x}\sin 2x$ are both solutions of the differential equation, the principle of superposition tells us that so is

$$y = Ce^{3x}\cos 2x + De^{3x}\sin 2x = e^{3x}(C\cos 2x + D\sin 2x),$$

for any values of C and D.

Exercise 20

Use the auxiliary equation to find the general solution of each of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} + 4\dfrac{dy}{dx} + 8y = 0$ (b) $\dfrac{d^2\theta}{dt^2} + 9\theta = 0$

We now summarise the method for solving these differential equations as a procedure.

Procedure 5 Solving homogeneous linear constant-coefficient second-order differential equations

The general solution of the homogeneous linear constant-coefficient second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0,$$

where a, b, c are (real) constants and $a \neq 0$, may be found as follows.

1. Write down the auxiliary equation

   $$a\lambda^2 + b\lambda + c = 0,$$

   and find its roots λ_1 and λ_2.

2. - If the auxiliary equation has two distinct real roots λ_1 and λ_2, the general solution of the differential equation is

     $$y = Ce^{\lambda_1 x} + De^{\lambda_2 x}.$$

   - If the auxiliary equation has two equal real roots $\lambda_1 = \lambda_2$, the general solution of the differential equation is

     $$y = (C + Dx)e^{\lambda_1 x}.$$

   - If the auxiliary equation has a pair of complex conjugate roots $\lambda_1 = \alpha + \beta i$ and $\lambda_2 = \alpha - \beta i$, the general solution of the differential equation is

     $$y = e^{\alpha x}(C\cos\beta x + D\sin\beta x).$$

   In each case, C and D are arbitrary constants.

It is worth noting that the three cases in Step 2 of Procedure 5 correspond to three different possibilities that arise when solving the characteristic equation $a\lambda^2 + b\lambda + c = 0$. These three different possibilities relate to the value of the **discriminant** $b^2 - 4ac$, where $b^2 - 4ac > 0$ corresponds to the first case, $b^2 - 4ac = 0$ to the second, and $b^2 - 4ac < 0$ to the third.

Exercise 21

Find the general solution of each of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} + 4y = 0$ (b) $u''(x) - 6u'(x) + 8u(x) = 0$

(c) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} = 0$ (d) $\dfrac{d^2y}{dx^2} - 2\dfrac{dy}{dx} + y = 0$

(e) $\dfrac{d^2y}{dx^2} - \omega^2 y = 0$, where ω is a real constant

(f) $\dfrac{d^2y}{dx^2} + 4\dfrac{dy}{dx} + 29y = 0$

Exercise 22

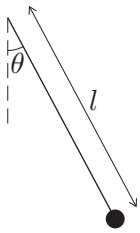


Figure 13 A pendulum

Small oscillations of the pendulum of a clock can be modelled by the differential equation

$$\ddot{\theta} = -\frac{g}{l}\theta,$$

where g is the magnitude of the acceleration due to gravity, l is the length of the pendulum, and θ is the angle that the pendulum makes with the vertical (see Figure 13).

Solve the differential equation to obtain an expression for θ in terms of g and l.

3 Inhomogeneous second-order differential equations

Section 2 was concerned with finding the general solution of homogeneous linear constant-coefficient second-order differential equations. This section concerns *inhomogeneous* linear constant-coefficient second-order differential equations, that is, equations of the form

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x), \tag{38}$$

where a, b, c are real constants, $a \neq 0$, and $f(x)$ is a given continuous real-valued function of x.

Subsection 3.1 gives the general method for constructing solutions of equation (38). Subsection 3.2 shows how to find an appropriate particular solution of the differential equation, for use in constructing the general solution, in cases where the function $f(x)$ takes one of a few particular forms. Subsection 3.3 deals with certain cases where complications can arise. Subsection 3.4 shows how to deal with cases where $f(x)$ is a combination of the functions discussed in Subsection 3.2.

3.1 General method of solution

The basic method used for finding the general solution of equation (38) depends on the principle of superposition (Theorem 2) and is illustrated in the following example.

Example 11

Show that $y = Ce^{-2x} + De^{-3x} + 2$ is a solution of the inhomogeneous differential equation

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 12, \tag{39}$$

for any values of the constants C and D.

Solution

We know from Exercise 18(a) that the homogeneous differential equation

$$\frac{d^2y}{dx^2} + 5\frac{dy}{dx} + 6y = 0 \tag{40}$$

has a general solution $y_c = Ce^{-2x} + De^{-3x}$, where C and D are arbitrary constants.

The notation y_c and y_p will be explained shortly.

Now consider the constant function $y_p = 2$. This is a particular solution of equation (39) since $d^2y_p/dx^2 = dy_p/dx = 0$ and $6y_p = 12$.

Therefore, by the principle of superposition (Theorem 2),

$$y = y_c + y_p = Ce^{-2x} + De^{-3x} + 2$$

is a solution of equation (39), for any values of C and D.

Equation (40) is an example of an *associated homogeneous equation* – that is, the homogeneous equation associated with the inhomogeneous equation (39) by making its right-hand side zero. The solutions y_c and y_p also have special names in this context: y_c, the general solution of the associated homogeneous equation (40), is called the *complementary function*, and y_p, a particular solution of the inhomogeneous equation (39), is called a *particular integral*.

(The term *particular integral* is used here, rather than the term particular solution used in some other texts, to distinguish it from the particular solution of equation (41) that satisfies given initial or boundary conditions (see Section 4).)

Let

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) \tag{41}$$

be an inhomogeneous linear constant-coefficient second-order differential equation.

Its **associated homogeneous equation** is

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0.$$

The general solution y_c of the associated homogeneous equation is known as the **complementary function** for the original inhomogeneous equation (41).

Any particular solution y_p of the original inhomogeneous equation (41) is referred to as a **particular integral** for that equation.

Later in this section we will show how to find particular integrals for a wide variety of equations. Before we do that, it is important that you realise the full significance of finding just one particular integral.

Exercise 23

Suppose that we have found two different particular integrals y_{p_1}, y_{p_2} for equation (41). Use the principle of superposition to show that the function $y_{p_1} - y_{p_2}$ is then a solution of the associated homogeneous equation.

That is, Section 2 enables us to find the complementary function.

The result of Exercise 23 shows the true significance of finding a particular integral. For if we do so, then since from Section 2 we know how to solve the associated homogeneous equation, we can find *all* particular integrals simply by adding the complementary function (which involves two arbitrary constants). We have the following important result.

Theorem 3

If y_c is the complementary function for an inhomogeneous linear constant-coefficient second-order differential equation, and y_p is a particular integral for that equation, then $y_c + y_p$ is the general solution of that equation.

Note that y_c, being the general solution of the associated homogeneous equation, will contain *two* arbitrary constants, whereas y_p, being a particular solution, will contain none.

Let us now see how the method based on the above theorem can be applied.

Example 12

Find the general solution of the differential equation

$$\frac{d^2y}{dx^2} + 9y = 9x + 9. \tag{42}$$

Solution

The associated homogeneous equation is

$$\frac{d^2y}{dx^2} + 9y = 0,$$

which has the general solution $y_c = C\cos 3x + D\sin 3x$, where C and D are arbitrary constants. This is the complementary function for equation (42).

See Exercise 20(b), although there different symbols were used for the variables.

A particular integral for equation (42) is $y_p = x + 1$.

You will see in the next subsection how to find such a particular integral.

This may be verified by differentiation and substitution: $y_p' = 1$ and $y_p'' = 0$, and substituting into the left-hand side of equation (42) gives

$$y_p'' + 9y_p = 0 + 9(x + 1) = 9x + 9,$$

which is the same as the right-hand side of equation (42), as required.

The general solution of equation (42) is therefore, by Theorem 3,

$$y = y_c + y_p = C\cos 3x + D\sin 3x + x + 1,$$

where C and D are arbitrary constants.

The method of Example 12 may be summarised as follows.

Procedure 6 Solving inhomogeneous linear constant-coefficient second-order differential equations

To find the general solution of the inhomogeneous linear constant-coefficient second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x),$$

carry out the following steps.

1. Find the complementary function y_c, that is, the general solution of the associated homogeneous equation

   $$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0,$$

   using Procedure 5.

   The reason why y_c is found first will become clear in Subsection 3.3.

2. Find a particular integral y_p.

The general solution is

$$y = y_c + y_p.$$

It is worth noting that *any* choice of particular integral in Procedure 6 gives the *same* general solution. Formulas obtained for the general solution may look different for different choices of particular integral, but they are in fact always equivalent. For example, in Example 12 the particular integral $y_p = x + 1$ was chosen, and the form of the general solution was obtained as $y = C\cos 3x + D\sin 3x + x + 1$. It would have been equally valid to have chosen as a particular integral $y_p = x + 1 + \sin 3x$. In that case, the form of the general solution would have been obtained as $y = C\cos 3x + D\sin 3x + x + 1 + \sin 3x$. This looks a little different, but it may be written in the form $y = C\cos 3x + (D+1)\sin 3x + x + 1$; and since C and D are arbitrary constants, this form of the general solution represents exactly the same family of solutions.

Exercise 24

For each of the following equations:

- write down its associated homogeneous equation and its complementary function y_c
- find a particular integral of the form $y_p = p$, where p is a constant
- write down the general solution.

See Exercise 21(a) and Example 8.

(a) $\dfrac{d^2y}{dx^2} + 4y = 8$ (b) $\dfrac{d^2y}{dx^2} - 3\dfrac{dy}{dx} + 2y = 6$

When using Procedure 6, the complementary function is found by using Procedure 5. However, the procedures for finding a particular integral are another matter. In Exercise 24, where the right-hand sides of the equations are constants, it was possible to find a particular integral almost 'by inspection'; but this method is generally inadequate. Fortunately, there exist procedures for finding a particular integral for equations involving wide classes of right-hand-side functions $f(x)$. The remainder of this section considers some of the simpler cases, where it is possible to determine the *form* of a particular integral by inspection, although some manipulation is required in order to determine the values of certain coefficients.

3.2 Finding a particular integral by the method of undetermined coefficients

In the previous subsection you saw that the inhomogeneous linear constant-coefficient second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x)$$

can be solved by first solving the associated homogeneous equation, using the methods of Section 2, and then finding a particular integral of the original equation, which depends on the function $f(x)$. This subsection and the next two subsections show you how to find a particular integral

when $f(x)$ is a polynomial, exponential or sinusoidal function, or a sum of such functions.

There exist procedures for finding a particular integral for fairly general types of continuous function $f(x)$, but these are not considered in this module.

You saw an example of the approach in Exercise 24. There the functions $f(x)$ were *constants* and you tried a *constant* function $y = p$ as a particular integral, substituting into the differential equation to find a suitable value for p. In general, we try a function of the same form as $f(x)$ as a particular integral, and substitute into the differential equation to find suitable values for its unknown coefficients. The function that we try is known as a **trial solution**, and the method is known as the method of **undetermined coefficients**.

The following examples illustrate the method. Bear in mind, though, that the method (and hence each example) finds only a particular integral for the differential equation; to find the general solution you would need to find the complementary function and combine it with the particular integral, according to Procedure 6.

A polynomial function

We consider $f(x) = m_n x^n + m_{n-1}x^{n-1} + \cdots + m_1 x + m_0$.

Let us first consider a case where $f(x)$ is a linear function (i.e. a polynomial of degree 1).

Example 13

Find a particular integral for

$$3\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 4x + 2.$$

Solution

We try a solution of the form

$$y = p_1 x + p_0,$$

where p_1 and p_0 are coefficients to be determined so that the differential equation is satisfied. To try this solution, we need the first and second derivatives of y:

$$\frac{dy}{dx} = p_1, \quad \frac{d^2y}{dx^2} = 0.$$

Substituting these into the left-hand side of the differential equation gives

$$3\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 3 \times 0 - 2p_1 + (p_1 x + p_0) = p_1 x + (p_0 - 2p_1).$$

Therefore for $y = p_1 x + p_0$ to be a solution of the differential equation, we require that

$$p_1 x + (p_0 - 2p_1) = 4x + 2 \quad \text{for all } x. \tag{43}$$

To find the two unknown coefficients p_1 and p_0, we compare the coefficients on both sides of equation (43). Comparing the terms in x gives $p_1 = 4$. Comparing the constant terms gives $p_0 - 2p_1 = 2$, so $p_0 = 2 + 2p_1 = 2 + 2 \times 4 = 10$.

Comparing coefficients works because two polynomials are equal if and only if all their corresponding coefficients are the same.

Therefore we have the particular integral

$$y_p = 4x + 10.$$

Check: If $y = 4x + 10$, then $dy/dx = 4$, $d^2y/dx^2 = 0$, and substituting into the left-hand side of the differential equation gives

$$3\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 3 \times 0 - 2 \times 4 + (4x + 10) = 4x + 2,$$

as required.

You will have noticed in Example 13 that substituting a *linear* trial solution $y = p_1x + p_0$ into the left-hand side of the differential equation resulted in a *linear* function, namely $p_1x + (p_0 - 2p_1)$, whose coefficients could be compared with those of the *linear* target function $4x + 2$ to obtain values for p_1 and p_0. This is really the key to the method. If the target function is linear, then choosing a linear trial solution ensures that substituting into the left-hand side of the differential equation results in a linear function whose coefficients can be compared with those of the target function. Similarly, as you will see below, if the target function belongs to one of certain other classes of functions, then choosing as a trial solution a *general* function from that class ensures that substitution into the left-hand side of the differential equation produces another function from the same class, whose coefficients can be compared with those of the target function, thus enabling values to be given to the coefficients of the trial solution. The method will work provided that *all* the derivatives of functions in the class are also in the class.

Exercise 25

Find particular integrals of the form $y = p_1x + p_0$ for each of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} - 2\dfrac{dy}{dx} + 2y = 2x + 3$

(b) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + y = 2x$

Try it and see what goes wrong.

Note that in Exercise 25(b), although $f(x)$ is just a multiple of x, it is not possible to find a solution of the form $y(x) = p_1x$. It is necessary for the trial solution to contain terms like those in $f(x)$ *and all its derivatives*, so in this case the trial solution must be of the form $y = p_1x + p_0$. So in general, even if $m_0 = 0$ in $f(x) = m_1x + m_0$, so that $f(x) = m_1x$, you should use a trial solution of the form $y = p_1x + p_0$.

You saw examples of this in Exercise 24.

However, if $f(x) = m_0$ is a constant function, then the trial solution need only be a constant function $y = p_0$.

In general, if $f(x) = m_n x^n + m_{n-1}x^{n-1} + \cdots + m_1 x + m_0$, where $m_n \neq 0$, then a trial solution of the form $y = p_n x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0$ should be used.

Exercise 26

Find a particular integral for

$$\ddot{y} - y = t^2.$$

An exponential function

We consider $f(x) = me^{kx}$.

Example 14

Find a particular integral for

$$\frac{d^2y}{dx^2} + 9y = 2e^{3x}.$$

Solution

We try a solution of the form

$$y = pe^{3x},$$

where p is a coefficient to be determined so that the differential equation is satisfied. Differentiating $y = pe^{3x}$ gives

$$\frac{dy}{dx} = 3pe^{3x}, \quad \frac{d^2y}{dx^2} = 9pe^{3x}.$$

Substituting these into the left-hand side of the differential equation gives

$$\frac{d^2y}{dx^2} + 9y = 9pe^{3x} + 9pe^{3x} = 18pe^{3x}.$$

Therefore for $y = pe^{3x}$ to be a solution of the differential equation, we require that $18pe^{3x} = 2e^{3x}$ for all x. Hence $p = \frac{1}{9}$, and

$$y_{\mathrm{p}} = \tfrac{1}{9}e^{3x}$$

is a particular integral for the given differential equation.

Since the derivative of e^{3x} is $3e^{3x}$, the exponent $(3x)$ appearing in $y(x)$ should be the same as that appearing in $f(x)$, and only the coefficient p is to be determined.

Exercise 27

Find a particular integral for

$$2\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = 2e^{-x}.$$

A sinusoidal function

This type of function is particularly important in many practical applications.

We consider $f(x) = m\cos\Omega x + n\sin\Omega x$.

Following on from earlier ideas, the trial solution must contain terms like those in $f(x)$ *and all its derivatives*; so even if $f(x)$ contains only a sine or only a cosine, the trial solution $y(x)$ must contain both a sine and a cosine. However, the value of the parameter Ω appearing in $y(x)$ should be the same as that appearing in $f(x)$.

Example 15

Find a particular integral for

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 2y = 10\sin 2x.$$

Solution

We try a solution of the form

$$y = p\cos 2x + q\sin 2x,$$

where p and q are coefficients to be determined so that the differential equation is satisfied. Differentiating y gives

$$\frac{dy}{dx} = -2p\sin 2x + 2q\cos 2x, \quad \frac{d^2y}{dx^2} = -4p\cos 2x - 4q\sin 2x.$$

Substituting these into the left-hand side of the differential equation gives

$$\begin{aligned}\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + 2y = {} & (-4p\cos 2x - 4q\sin 2x) \\ & + 2(-2p\sin 2x + 2q\cos 2x) \\ & + 2(p\cos 2x + q\sin 2x) \\ = {} & (-2p + 4q)\cos 2x + (-4p - 2q)\sin 2x.\end{aligned}$$

Therefore for $y = p\cos 2x + q\sin 2x$ to be a solution of the differential equation, we require that

$$(-2p + 4q)\cos 2x + (-4p - 2q)\sin 2x = 10\sin 2x \quad \text{for all } x. \tag{44}$$

Comparing coefficients works because the cosine and sine functions are *linearly independent*: if $a\sin rx + b\cos rx = 0$ for all x, then $a = b = 0$.

To find p and q, we compare the coefficients of $\cos 2x$ and $\sin 2x$ on both sides of equation (44). Comparing $\cos 2x$ terms gives $-2p + 4q = 0$, and comparing $\sin 2x$ terms gives $-4p - 2q = 10$. Solving these simultaneous equations gives $p = -2$, $q = -1$. Hence

$$y_{\mathrm{p}} = -2\cos 2x - \sin 2x$$

is a particular integral for the given differential equation.

Exercise 28

Find a particular integral for

$$\frac{d^2y}{dt^2} - \frac{dy}{dt} = \cos 3t + \sin 3t.$$

Method of undetermined coefficients

The following procedure summarises the results of this subsection.

Procedure 7 Method of undetermined coefficients

To find a particular integral for the inhomogeneous linear constant-coefficient second-order differential equation

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x),$$

use a trial solution $y(x)$ that has a form similar to that of $f(x)$. For simple forms of $f(x)$, the following table gives the appropriate form of trial solution. (In the table, m_i, p_i $(i = 1, \ldots, n)$, m, p, q and Ω are all constants.)

Note that there are exceptional cases where these trial solutions do not work; see Subsection 3.3.

$f(x)$	Trial solution $y(x)$
$m_n x^n + m_{n-1}x^{n-1} + \cdots + m_1 x + m_0$	$p_n x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0$
me^{kx}	pe^{kx}
$m\cos\Omega x + n\sin\Omega x$	$p\cos\Omega x + q\sin\Omega x$

To determine the coefficient(s) in $y(x)$, differentiate $y(x)$ twice, substitute into the left-hand side of the differential equation, and equate coefficients of corresponding terms.

Exercise 29

What form of trial solution should you use in order to find a particular integral for each of the following differential equations?

(a) $\dfrac{d^2y}{dx^2} - y = e^{2x}$

(b) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} - 4y = \sin 4x$

Exercise 30

Use Procedures 6 and 7 to find the *general* solution of each of the following differential equations.

The roots of the auxiliary equation are $-1 \pm i$ for part (a), and 0 and -3 for part (b).

(a) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + 2y = 4$

(b) $\dfrac{d^2\theta}{dt^2} + 3\dfrac{d\theta}{dt} = 13\cos 2t$

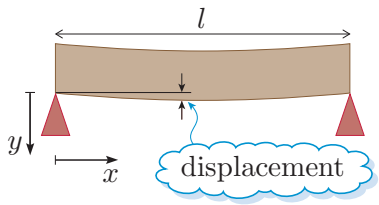


Figure 14 A horizontal beam resting on supports

Exercise 31

A long horizontal rectangular beam of length l rests on rigid supports at each end. It is important in civil engineering to determine how much such a beam 'sags'. A simple model of this 'sag', or vertical displacement y, is the differential equation

$$Ry'' - Sy + \tfrac{1}{2}Q(l - x)x = 0,$$

where R, S and Q are constants related to the structure of the beam, and x is the distance from one end of the beam (as illustrated in Figure 14).

Find the *general* solution of the differential equation in the case where R, S and Q are all equal to 1.

In Subsection 3.4 you will see how the principle of superposition can be used in combination with Procedure 7 to solve differential equations whose right-hand-side functions $f(x)$ are sums of polynomial, exponential and sinusoidal functions. But first let us look at some exceptional cases for which Procedure 7 does not work and needs to be adapted.

3.3 Exceptional cases

There are some exceptional cases for which Procedure 7 fails. The aim of this subsection is to indicate when such difficulties arise, and how a particular integral may be found in those circumstances. Let us begin with an example.

Example 16

Find a particular integral for

$$\frac{d^2y}{dx^2} - 4y = 2e^{2x}.$$

Solution

Using Procedure 7, let us try $y = pe^{2x}$. Differentiating this gives

$$\frac{dy}{dx} = 2pe^{2x}, \quad \frac{d^2y}{dx^2} = 4pe^{2x}.$$

Substituting these into the left-hand side of the differential equation gives

$$\frac{d^2y}{dx^2} - 4y = 4pe^{2x} - 4pe^{2x} = 0.$$

So there is no value of p that gives a particular integral of the form $y = pe^{2x}$.

The trouble is that the complementary function, that is, the general solution of the associated homogeneous equation

$$\frac{d^2y}{dx^2} - 4y = 0,$$

is $y = Ce^{-2x} + De^{2x}$, where C and D are arbitrary constants, thus the trial solution is a solution of the associated homogeneous equation (with

See Exercise 18(c).

$C = 0$, $D = p$). Hence on substituting the trial solution $y = pe^{2x}$ into the inhomogeneous equation, the left-hand side is zero for any value of p, so it cannot be equal to the non-zero right-hand side.

In such circumstances, the difficulty can generally be overcome by multiplying the trial solution suggested in Procedure 7 by x. Thus in this case, the trial solution should be modified to take the form $y = pxe^{2x}$. Differentiating this gives

$$\frac{dy}{dx} = pe^{2x} + 2pxe^{2x} = p(1 + 2x)e^{2x},$$
$$\frac{d^2y}{dx^2} = 2pe^{2x} + 2p(1 + 2x)e^{2x} = 4p(1 + x)e^{2x}.$$

There is an analogy here with the case of the homogeneous differential equation when the characteristic equation has equal roots; in that case, when $e^{\lambda x}$ is one solution of the equation, another solution is given by $xe^{\lambda x}$.

Substituting these into the left-hand side of the differential equation gives

$$\frac{d^2y}{dx^2} - 4y = 4p(1 + x)e^{2x} - 4pxe^{2x} = 4pe^{2x}.$$

Therefore $y = pxe^{2x}$ is a solution of the differential equation provided that $4pe^{2x} = 2e^{2x}$ for all x. Hence $p = \frac{1}{2}$, and

$$y_\mathrm{p} = \tfrac{1}{2}xe^{2x}$$

is a particular integral for the given differential equation.

The problem with the trial solution being a solution of the associated homogeneous equation can occur irrespective of the form of the trial solution (i.e. polynomial, exponential or sinusoidal), but in most cases it can be overcome by multiplying the trial solution suggested in Procedure 7 by x. When using Procedure 7, you should check whether the proposed trial solution is a solution of the associated homogeneous equation, and if so try multiplying it by x. (This is why it is important to find y_c before y_p in Procedure 6.)

Exercise 32

Find a particular integral for each of the following differential equations.

(a) $\dfrac{d^2y}{dx^2} - 3\dfrac{dy}{dx} + 2y = 4e^x$ (b) $2\dfrac{d^2y}{dx^2} + 3\dfrac{dy}{dx} = 1$

The complementary functions are given in the solutions to Example 8 and Exercise 18(b).

Exercise 33

The motion of a marble dropped from the Clifton Suspension Bridge into the River Avon can be modelled by the differential equation

$$m\ddot{x} + r\dot{x} - mg = 0,$$

where m is the mass of the marble, r is a constant related to air resistance, g is the magnitude of the acceleration due to gravity, and x is the vertical distance from the point of dropping (as shown in Figure 15).

Find an expression for x in terms of the time t.

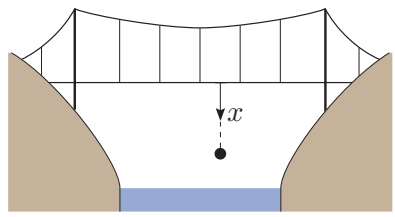


Figure 15 A marble dropped from the Clifton Suspension Bridge

---

We have seen that Procedure 7 fails if the trial solution is a solution of the associated homogeneous differential equation; in such cases we multiply the suggested trial solution by the independent variable and use this as the trial solution. Another situation in which it is necessary to multiply the suggested trial solution by the independent variable is illustrated in the following example.

Example 17

Find a particular integral for

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} = 2x + 2.$$

Solution

Using Procedure 7, let us try $y = p_1x + p_0$. Differentiating this gives

$$\frac{dy}{dx} = p_1, \quad \frac{d^2y}{dx^2} = 0.$$

Substituting these into the left-hand side of the differential equation gives

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} = 2p_1.$$

But there is no value of p_1 that satisfies $2p_1 = 2x + 2$ for *all* x.

This example again illustrates why it is better to find the complementary function before looking for a particular integral.

The problem this time is that (from Exercise 21(c)) the complementary function is $y = C + De^{-2x}$, where C and D are arbitrary constants, so the p_0 part of the trial solution is a solution of the associated homogeneous equation (with $C = p_0$, $D = 0$). Hence on substituting the trial solution $y = p_1x + p_0$ into the inhomogeneous equation, the p_0 part disappears, and the trial solution effectively reduces to $y = p_1x$. The result in this case is that after substituting the trial solution and its derivatives into the left-hand side of the equation, there are not enough terms on the left-hand side to compare with the terms in the right-hand-side function.

Again, the difficulty can be overcome by multiplying the trial solution suggested by Procedure 7 by x, to give $y = p_1x^2 + p_0x$. Differentiating this gives

$$\frac{dy}{dx} = 2p_1x + p_0, \quad \frac{d^2y}{dx^2} = 2p_1.$$

Substituting these into the left-hand side of the differential equation gives

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} = 2p_1 + 2(2p_1x + p_0) = 4p_1x + (2p_1 + 2p_0).$$

Therefore $y = p_1x^2 + p_0x$ is a solution of the differential equation provided that $4p_1x + (2p_1 + 2p_0) = 2x + 2$ for all x. This gives $p_1 = \frac{1}{2}$, $p_0 = \frac{1}{2}$, so

$$y_\mathrm{p} = \tfrac{1}{2}(x^2 + x)$$

is a particular integral for the given differential equation.

To summarise, Procedure 7 will fail if *all* or *part* of the suggested trial solution is a solution of the associated homogeneous equation. In such cases, a particular integral can usually be found by multiplying the trial solution by the independent variable.

However, it may sometimes be necessary to multiply the trial function by the independent variable more than once, as explained below in Procedure 8 and Exercise 34.

Procedure 8 Exceptional cases in the method of undetermined coefficients

Suppose that you try using the method of undetermined coefficients (described in Procedure 7) for finding a particular integral for an inhomogeneous linear constant-coefficient second-order differential equation.

If this fails because all or part of the trial solution is a solution of the associated homogeneous equation, then try multiplying the trial solution by the independent variable.

If all or part of the resulting trial solution is still a solution of the associated homogeneous equation, then try multiplying by the independent variable again.

Exercise 34

Find a particular integral for

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = e^x.$$

You found the complementary function in Exercise 21(d).

3.4 Combining cases

You have seen how to find a particular integral when the right-hand-side function $f(x)$ is polynomial, exponential or sinusoidal. In this subsection you will see how to find a particular integral when $f(x)$ is a combination of such functions, by using the principle of superposition.

Example 18

Find a particular integral for

$$\frac{d^2y}{dx^2} + 9y = 2e^{3x} + 18x + 18. \tag{45}$$

Solution

In Example 14 you saw that $y_p = \frac{1}{9}e^{3x}$ is a particular integral for

$$\frac{d^2y}{dx^2} + 9y = 2e^{3x},$$

and in Example 12 you saw that $y_p = x + 1$ is a particular integral for

$$\frac{d^2y}{dx^2} + 9y = 9x + 9.$$

Therefore, by the principle of superposition, a particular integral for equation (45) is

$$y_p = \tfrac{1}{9}e^{3x} + 2 \times (x + 1) = \tfrac{1}{9}e^{3x} + 2x + 2.$$

The approach of Example 18 is to find particular integrals for differential equations involving each part of $f(x)$ separately, and then to use the principle of superposition to combine the two.

Exercise 35

Find particular integrals for each of the following differential equations.

See Exercise 34.

(a) $\dfrac{d^2y}{dx^2} - 2\dfrac{dy}{dx} + y = 4e^x - 3e^{2x}$

(b) $2\dfrac{d^2x}{dt^2} + 3\dfrac{dx}{dt} + 2x = 12\cos 2t + 10$

Exercise 36

Find the *general* solution of each of the following differential equations.

You will find some help for parts (a), (d), (e) and (f) in Exercises 18 and 21, and Example 8.

(a) $\dfrac{d^2\theta}{dt^2} + 4\theta = 2t$

(b) $u''(t) + 4u'(t) + 5u(t) = 5$

(c) $3\dfrac{d^2Y}{dx^2} - 2\dfrac{dY}{dx} - Y = e^{2x} + 3$

(d) $\dfrac{d^2y}{dx^2} - 4y = e^{-2x}$

(e) $\dfrac{d^2y}{dx^2} + 4y = \sin 2x + 3x$

(f) $\dfrac{d^2y}{dx^2} - 3\dfrac{dy}{dx} + 2y = 2e^x - 5e^{2x}$

---

4 Initial conditions and boundary conditions

In Section 3 you saw how to find the general solution of an inhomogeneous linear constant-coefficient second-order differential equation as a combination of a complementary function and a particular integral. In practice, however, we usually want a *particular* solution that satisfies certain additional conditions. Recall that a particular solution is one that does not involve arbitrary constants. In Section 1 you saw how one additional condition (called an initial condition) was needed to find a value for the single arbitrary constant in the general solution of a first-order differential equation in order to obtain a particular solution. In the case of second-order differential equations, in order to obtain a particular solution, *two* additional conditions are needed to obtain values for the *two* arbitrary constants in the general solution.

There are two types of additional conditions for second-order differential equations: *initial conditions* and *boundary conditions*. Problems involving such conditions are called *initial-value problems* and *boundary-value problems*, respectively, and are discussed in Subsections 4.1 and 4.2.

4.1 Initial-value problems

For a *first-order* differential equation, an initial condition consists of specifying the value of the dependent variable ($y = y_0$, say) at a given value of the independent variable ($x = x_0$), and is often written in the form $y(x_0) = y_0$.

See Section 1.

One fairly obvious way of specifying *two* additional conditions for a *second-order* differential equation is to give the values of both the dependent variable ($y = y_0$) and its derivative ($dy/dx = z_0$) for the *same* given value of the independent variable ($x = x_0$).

There are many examples of such a pair of initial conditions occurring naturally as part of a problem. One example is provided by the marble being dropped from the Clifton Suspension Bridge in Exercise 33. In that example, x is the vertical distance from the point of dropping. The obvious choice of origin for the time t is the time at which the marble is dropped. Therefore a naturally occurring pair of initial conditions is that at time $t = 0$, we know both the position $x = 0$ and the speed $\dot{x} = 0$ (since the marble is dropped, i.e. is released with zero initial velocity). Another example is provided by the clock pendulum in Exercise 22. In this example, when the pendulum changes direction, its rate of change of angle θ is momentarily zero; also, when it changes direction, it makes its greatest angle θ_0 with the vertical (see Figure 16). Therefore if we measure time t from the moment when the pendulum changes direction, we have the initial conditions $\theta = \theta_0$ and $\dot{\theta} = 0$ when $t = 0$.

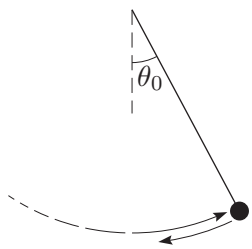


Figure 16 A pendulum changing direction

Initial conditions associated with a second-order differential equation with dependent variable y and independent variable x specify that y and dy/dx take values y_0 and z_0, respectively, when x takes the value x_0. These conditions can be written as

$$y = y_0 \text{ and } \frac{dy}{dx} = z_0 \text{ when } x = x_0$$

or as

$$y(x_0) = y_0, \quad y'(x_0) = z_0.$$

The numbers x_0, y_0 and z_0 are often referred to as **initial values**.

The combination of a second-order differential equation and initial conditions is called an **initial-value problem**.

Let us now see how initial conditions can be used to determine values for the two arbitrary constants and hence find a particular solution.

Example 19

Find the particular solution of the differential equation

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 0$$

that satisfies the initial conditions $y = 0$ and $dy/dx = 1$ when $x = 0$.

Solution

From Example 8 we know that the general solution is

$$y = Ce^x + De^{2x}, \tag{46}$$

where C and D are arbitrary constants. One of the initial conditions involves the derivative of the solution, so we need to obtain that derivative:

$$\frac{dy}{dx} = Ce^x + 2De^{2x}. \tag{47}$$

The initial conditions state that $y(0) = 0$, $y'(0) = 1$. Substituting $x = 0$, $y = 0$ into equation (46) gives

$$0 = Ce^0 + De^0 = C + D,$$

while substituting $x = 0$, $dy/dx = 1$ into equation (47) gives

$$1 = Ce^0 + 2De^0 = C + 2D.$$

Note that when you check a particular solution, you should check that it satisfies the initial conditions as well as the differential equation.

Solving these equations gives $C = -1$, $D = 1$, so the required particular solution is

$$y = -e^x + e^{2x}.$$

Exercise 37

Find solutions to the following initial-value problems.

(a) $u''(t) + 9u(t) = 0, \quad u\left(\frac{\pi}{2}\right) = 0,\ u'\left(\frac{\pi}{2}\right) = 1.$ See Exercise 20(b).

(b) $\dfrac{d^2y}{dx^2} - 3\dfrac{dy}{dx} + 2y = 4e^x, \quad y(0) = 4,\ y'(0) = 2.$ See Exercise 32(a).

(c) $\dfrac{d^2y}{dx^2} - 2\dfrac{dy}{dx} + y = 4e^x - 3e^{2x}, \quad y(0) = 4,\ y'(0) = -1.$ See Exercises 21(d) and 35(a).

You saw in Subsection 1.4 that an initial-value problem involving a linear first-order differential equation has a unique solution under certain circumstances. (Such circumstances hold for nearly every such initial-value problem that you are likely to come across in practice.) The same is true of initial-value problems involving a linear constant-coefficient second-order differential equation, as the following theorem makes clear.

Theorem 4

The initial-value problem

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x), \quad y(x_0) = y_0,\ y'(x_0) = z_0,$$

where a, b, c are real constants with $a \neq 0$, and $f(x)$ is a given continuous real-valued function on an interval (r, s), with $x_0 \in (r, s)$, has a unique solution on that interval.

Note that one consequence of this theorem is that if the differential equation is *homogeneous* and the initial conditions are of the form $y(x_0) = 0$ and $y'(x_0) = 0$, then the unique solution must be the zero function $y = 0$, since it satisfies the differential equation and the initial conditions.

4.2 Boundary-value problems

The two conditions in an initial-value problem (the value of the dependent variable y and its derivative dy/dx) both relate to the same value of x. However, the two conditions that are required to determine values for the arbitrary constants need not relate to the same value of x. We could have one condition for $x = x_0$ and another for $x = x_1$, say. For example, consider again the 'sagging' beam from Exercise 31. Two known conditions on this beam are its zero displacements at the ends of the beam, where it rests on the rigid supports: that is, its *boundary* conditions are $y(0) = 0$ and $y(l) = 0$ (where l is the length of the beam). This pair of boundary conditions gives the value of y at two different points, but in general each boundary condition could specify the value of either y or dy/dx (or even a relationship between them).

Boundary conditions associated with a second-order differential equation with dependent variable y and independent variable x specify that y or dy/dx (or some combination of the two) takes values y_0 and y_1 at two different values x_0 and x_1, respectively, of x. The numbers x_0, x_1, y_0 and y_1 are often referred to as **boundary values**.

The combination of a second-order differential equation and boundary conditions is called a **boundary-value problem**.

The conditions are referred to as 'boundary' conditions because, as in the beam example, they often relate to conditions at the endpoints x_0 and x_1 of an interval $[x_0, x_1]$ on which we are interested in exploring the differential equation.

Let us now see how boundary conditions can be used to determine values for the two arbitrary constants and hence find a particular solution.

Example 20

Find the particular solution of the differential equation

$$\frac{d^2y}{dx^2} + 9y = 0$$

that satisfies the boundary conditions $y = 0$ when $x = 0$ and $dy/dx = 1$ when $x = \frac{\pi}{3}$.

Solution

From Exercise 20(b), the general solution is

$$y = C\cos 3x + D\sin 3x, \tag{48}$$

where C and D are arbitrary constants.

One of the boundary conditions involves the derivative of the solution, so we need to obtain that derivative:

$$\frac{dy}{dx} = -3C\sin 3x + 3D\cos 3x. \tag{49}$$

The boundary conditions state that $y(0) = 0$, $y'\left(\frac{\pi}{3}\right) = 1$. Substituting $x = 0$, $y = 0$ into equation (48) gives

$$0 = C\cos 0 + D\sin 0 = C,$$

so $C = 0$. Substituting $x = \frac{\pi}{3}$, $y' = 1$ and $C = 0$ into equation (49) gives

$$1 = 3D\cos\pi = -3D.$$

Therefore $C = 0$, $D = -\frac{1}{3}$, so the required particular solution is

$$y = -\tfrac{1}{3}\sin(3x).$$

Exercise 38

Solve the boundary-value problem

$$\frac{d^2y}{dx^2} - 3\frac{dy}{dx} + 2y = 4e^x, \quad y'(0) = 2,\ y(1) = 0.$$

See Exercise 32(a).

Exercise 39

Use the differential equation of Exercise 31, with $R = S = Q = 1$, namely

$$y'' - y + \tfrac{1}{2}(l - x)x = 0,$$

to determine the vertical displacement at the centre of a beam of length 2 metres resting on rigid supports at its ends.

Unlike the case of initial-value problems, boundary-value problems may not have solutions even when the differential equation is linear and constant-coefficient with a continuous real-valued right-hand-side function, as the following example illustrates.

Example 21

Try to solve the boundary-value problem

$$\frac{d^2y}{dx^2} + 4y = 0, \quad y(0) = 0,\ y\left(\tfrac{\pi}{2}\right) = 1.$$

Solution

From Exercise 21(a), the general solution is

$$y = C\cos 2x + D\sin 2x,$$

where C and D are arbitrary constants.

The boundary conditions state that $y(0) = 0$, $y\left(\frac{\pi}{2}\right) = 1$. Substituting each of these into the general solution in turn gives

$$0 = C\cos 0 + D\sin 0 = C,$$
$$1 = C\cos\pi + D\sin\pi = -C.$$

There is no solution for which $C = 0$ and $C = -1$, so there is no solution of the differential equation that satisfies the given boundary conditions.

Fortunately it is rare for a boundary-value problem that models a real-life situation to have no solution (and in such cases it is usually possible to reformulate the model to overcome the difficulty).

Not only is it possible for boundary-value problems to have no solution, but it is also possible for them to have solutions that are not unique, as the following example illustrates.

Example 22

Solve the boundary-value problem

$$\frac{d^2y}{dx^2} + 4\frac{dy}{dx} + 5y = 5, \quad y(0) = 1,\ y(\pi) = 1.$$

Solution

From Exercise 36(b), the general solution is

$$y = e^{-2x}(C\cos x + D\sin x) + 1,$$

where C and D are arbitrary constants.

The boundary conditions state that $y(0) = 1$, $y(\pi) = 1$. Substituting each of these into the general solution in turn gives

$$1 = e^0(C\cos 0 + D\sin 0) + 1 = C + 1,$$
$$1 = e^{-2\pi}(C\cos\pi + D\sin\pi) + 1 = -Ce^{-2\pi} + 1.$$

Both of these equations reduce to $C = 0$, but D can take any value, so any solution of the form

$$y = De^{-2x}\sin x + 1$$

satisfies the differential equation and the boundary conditions.

In Example 22, there is no unique solution of the differential equation that satisfies the given boundary conditions, but instead there is an infinite family of possible solutions.

Finally, a word of reassurance: most of the boundary-value problems that you will come across in this module will have a unique solution.

The final exercises of the unit test your understanding of the whole of this section.

Exercise 40

You found the general solution of the differential equation in Exercise 21(a).

For each of the following problems, identify the conditions as either initial conditions or boundary conditions, and solve the problem.

(a) $u''(x) + 4u(x) = 0, \quad u(0) = 1,\ u'(0) = 0.$

(b) $u''(x) + 4u(x) = 0, \quad u(0) = 0,\ u\left(\frac{\pi}{2}\right) = 0.$

(c) $u''(x) + 4u(x) = 0, \quad u\left(\frac{\pi}{2}\right) = 0,\ u'\left(\frac{\pi}{2}\right) = 0.$

(d) $u''(x) + 4u(x) = 0, \quad u(-\pi) = 1,\ u\left(\frac{\pi}{4}\right) = 2.$

(e) $u''(x) + 4u(x) = 0, \quad u'(0) = 0,\ u'\left(\frac{\pi}{4}\right) = 1.$

Exercise 41

Find the solution (if any) to each of the following problems.

(a) $u''(t) + 4u'(t) + 5u(t) = 0, \quad u(0) = 0,\ u'(0) = 2.$

The roots of the auxiliary equation are $-2 \pm i$.

(b) $\dfrac{d^2y}{dx^2} + 2\dfrac{dy}{dx} + 2y = 0, \quad$ where $y = 0$ and $\dfrac{dy}{dx} = 0$ when $x = 0.$

Roots are $-1 \pm i$.

(c) $\ddot{x} + 9x = 3(1 - \pi t), \quad x(0) = \frac{1}{3},\ \dot{x}\left(\frac{\pi}{3}\right) = 0.$

Roots are $\pm 3i$.

Learning outcomes

After studying this unit, you should be able to:

- understand and use the basic terminology relating to differential equations and their solutions
- check by substitution whether a given function is a solution of a given first-order differential equation or initial-value problem
- find from the general solution of a first-order differential equation the particular solution that satisfies a given initial condition
- appreciate the difficulties with domains and image sets for the solution of some differential equations
- deduce the qualitative behaviour of solutions from consideration of a first-order differential equation itself, as visualised from its direction field
- set up the formulas required by Euler's method for solving an initial-value problem, and carry out a few steps of the method
- recognise when a first-order differential equation is soluble by direct integration, and carry out that integration when appropriate, in simple cases
- recognise when a first-order differential equation is separable, and apply the method of separation of variables in simple cases
- recognise when a first-order differential equation is linear, and solve such an equation by the integrating factor method in simple cases
- understand and use the terminology relating to linear constant-coefficient second-order differential equations
- understand the key role of the principle of superposition in the solution of linear constant-coefficient second-order differential equations
- obtain the general solution of a homogeneous linear constant-coefficient second-order differential equation using the solutions of its auxiliary equation

- use the method of undetermined coefficients to find a particular integral for an inhomogeneous linear constant-coefficient second-order differential equation with certain simple forms of right-hand-side function
- obtain the general solution of an inhomogeneous linear constant-coefficient second-order differential equation by combining its complementary function and a particular integral
- use the general solution together with a pair of initial or boundary conditions to obtain, when possible, a particular solution of a linear constant-coefficient second-order differential equation.

Solutions to exercises

Solution to Exercise 1

We have

$$r(P) = k\left(1 - \frac{P}{M}\right),$$

so we need to solve the pair of simultaneous equations

$$k\left(1 - \frac{10}{M}\right) = 1, \quad k\left(1 - \frac{10\,000}{M}\right) = 0.$$

From the second equation, since $k > 0$, we see immediately that $M = 10\,000$. Substituting in the first equation leads to

$$k\frac{999}{1000} = 1, \quad \text{so} \quad k = \frac{1000}{999}.$$

Solution to Exercise 2

(a) In each case we need to show that the given function satisfies the differential equation, that is, it gives the same expression for either side of the equation.

If $y = 2e^x - (x^2 + 2x + 2)$, then differentiating y gives

$$\frac{dy}{dx} = 2e^x - 2x - 2,$$

and substituting the expression for y into the right-hand side of the given differential equation gives

$$y + x^2 = 2e^x - (x^2 + 2x + 2) + x^2 = 2e^x - 2x - 2,$$

as required.

(b) If $y = \tan x + \sec x$, then

$$\frac{dy}{dx} = \sec^2 x + \tan x \sec x,$$

and substituting the expression for y into the right-hand side of the given differential equation gives

$$\begin{aligned} y\tan x + 1 &= (\tan x + \sec x)\tan x + 1 \\ &= (\tan^2 x + 1) + \sec x \tan x \\ &= \sec^2 x + \tan x \sec x, \end{aligned}$$

as required.

(c) If $y = t + Ce^{-t}$, then

$$\dot{y} = \frac{dy}{dt} = 1 - Ce^{-t},$$

and substituting the expression for y into the right-hand side of the given differential equation gives

$$-y + t + 1 = -(t + Ce^{-t}) + t + 1 = 1 - Ce^{-t},$$

as required.

Solution to Exercise 3

(a) In each case we need to show that the given function satisfies the differential equation, that is, it gives the same expression for either side of the equation.

If $y = C - \frac{1}{3}e^{-3x}$, then

$$\frac{dy}{dx} = e^{-3x},$$

as required.

(b) If

$$P = \frac{CMe^{kt}}{1 + Ce^{kt}},$$

Rules such as the quotient rule can be found in the Handbook.

then, using the quotient rule for differentiation,

$$\begin{aligned}\frac{dP}{dt} &= \frac{(1 + Ce^{kt})(CMke^{kt}) - (CMe^{kt})(Cke^{kt})}{(1 + Ce^{kt})^2} \\ &= k\left(\frac{CMe^{kt}}{1 + Ce^{kt}}\right)\left(\frac{1 + Ce^{kt} - Ce^{kt}}{1 + Ce^{kt}}\right) \\ &= k\left(\frac{CMe^{kt}}{1 + Ce^{kt}}\right)\left(1 - \frac{Ce^{kt}}{1 + Ce^{kt}}\right) \\ &= kP\left(1 - \frac{P}{M}\right),\end{aligned}$$

as required.

Solution to Exercise 4

(a) From Exercise 3(b) and using $k = 0.15$ and $M = 10$, we know that

$$P = \frac{CMe^{kt}}{1 + Ce^{kt}} = \frac{10Ce^{0.15t}}{1 + Ce^{0.15t}}$$

is a solution of the differential equation, where C is a constant. The initial condition $P(0) = 1$ then implies

$$1 = \frac{10C}{1 + C}, \quad \text{so} \quad C = \tfrac{1}{9}.$$

A particular solution is therefore

$$P = \frac{\frac{10}{9}e^{0.15t}}{1 + \frac{1}{9}e^{0.15t}} = \frac{10e^{0.15t}}{9 + e^{0.15t}}.$$

(b) Dividing top and bottom by $e^{0.15t}$, we see that

$$P = \frac{10}{9e^{-0.15t} + 1}.$$

For large values of t, the exponential term on the bottom will be very small. The result is that P will approach the value 10 in the long term.

Solution to Exercise 5

(a) In each case, we apply direct integration to find the general solution, and C is an arbitrary constant.

The differential equation $dy/dx = 6x$ has general solution

$$y = \int 6x\,dx = 3x^2 + C.$$

From the initial condition $y(1) = 5$, we have $5 = 3 + C$, so $C = 2$. The solution to the initial-value problem is therefore

$$y = 3x^2 + 2.$$

(b) The differential equation $dv/du = e^{4u}$ has general solution

$$v = \int e^{4u}\,du = \tfrac{1}{4}e^{4u} + C.$$

From the initial condition $v(0) = 2$, we have $2 = \frac{1}{4} + C$, so $C = \frac{7}{4}$. The solution to the initial-value problem is therefore

$$v = \tfrac{1}{4}e^{4u} + \tfrac{7}{4}.$$

(c) The differential equation $\dot{y} = t/(1 + t^2)$ has general solution

$$y = \int \frac{t}{1+t^2}\,dt.$$

Using the hint, we make the substitution $u = 1 + t^2$, for which $du/dt = 2t$. This gives

$$\begin{aligned}\int \frac{t}{1+t^2}\,dt &= \tfrac{1}{2}\int \frac{1}{1+t^2}(2t)\,dt \\ &= \tfrac{1}{2}\int \frac{1}{u}\,du \\ &= \tfrac{1}{2}\ln u + C \quad (\text{since } u = 1 + t^2 > 0) \\ &= \tfrac{1}{2}\ln(1+t^2) + C.\end{aligned}$$

The $\frac{1}{2}$ is included outside the integral to match the 2 inside the integral that enables the substitution to be made.

The general solution of the differential equation is therefore

$$y = \tfrac{1}{2}\ln(1+t^2) + C.$$

The initial condition $y(0) = 2$ gives $C = 2$, so the particular solution is

$$y = \tfrac{1}{2}\ln(1+t^2) + 2.$$

Solution to Exercise 6

(a) Each of the differential equations is soluble by direct integration.

The general solution of $dy/du = 1/(u - a)$, where $u \neq a$, is given by

$$y = \int \frac{1}{u-a}\,du.$$

Using equation (8), integration produces the general solution

$$y = \ln|u - a| + C,$$

where C is an arbitrary constant.

(b) To verify the equation in the hint, taking a common denominator, we have

$$\frac{1}{x}+\frac{a}{1-ax}=\frac{(1-ax)+ax}{x(1-ax)}=\frac{1}{x(1-ax)}.$$

Then the general solution of $dy/dx = 1/(x(1-ax))$, where $x \neq 0$, $x \neq 1/a$, is given by

$$\begin{aligned} y=\int \frac{1}{x(1-ax)}\,dx &= \int \frac{1}{x}\,dx+\int \frac{a}{1-ax}\,dx \\ &= \ln|x|-\ln|1-ax|+C \\ &= \ln\left|\frac{x}{1-ax}\right|+C, \end{aligned}$$

where C is an arbitrary constant. This can also be written as

$$y=C-\ln\left|\frac{1-ax}{x}\right|=C-\ln\left|\frac{1}{x}-a\right|.$$

Solution to Exercise 7

(a) The differential equation is $dm/dt = -\lambda m$, where $m > 0$. Following Procedure 2, we obtain

$$\int \frac{1}{m}\,dm=\int(-\lambda)\,dt,$$

and since $m > 0$, integration produces

$$\ln m=-\lambda t+C,$$

where C is an arbitrary constant. On solving this equation for m, by taking the exponential of both sides, we obtain

$$m=e^{-\lambda t+C}=e^{C}e^{-\lambda t}=Be^{-\lambda t},$$

where $B = e^C$ is a positive (since $e^C > 0$ for all C) but otherwise arbitrary constant. The general solution is therefore

$$m=Be^{-\lambda t},$$

where B is a positive but otherwise arbitrary constant.

(b) The initial condition is $m(0) = m_0$, from which we have $m_0 = Be^0$, so $B = m_0$. The required particular solution is therefore

$$m=m_0e^{-\lambda t}.$$

Solution to Exercise 8

(a) The differential equation is

$$\frac{dy}{dx}=\frac{y-1}{x}, \quad \text{where } x>0.$$

In order to apply the separation of variables method, we need to exclude the case where $y = 1$.

So for $y \neq 1$, on applying Procedure 2 we have

$$\int \frac{1}{y-1}\,dy = \int \frac{1}{x}\,dx.$$

Since $x > 0$, for $y \neq 1$ (so that $y - 1 \neq 0$), integration produces

$$\ln|y-1| = \ln x + C,$$

where C is an arbitrary constant. Writing $C = \ln B$ ($B > 0$) and using properties of the logarithm function, the right-hand side becomes $\ln x + \ln B = \ln(Bx)$, so we can deduce that

$$|y-1| = Bx.$$

So $y - 1 = Bx$ if $y > 1$, and $y - 1 = -Bx$ if $y < 1$. We thus have

$$y = 1 + Ax,$$

where the sign can be absorbed into A, which is a non-zero but otherwise arbitrary constant.

Examination of the differential equation shows that $A = 0$ also gives a solution (the constant function $y = 1$).

The general solution is therefore

$$y = 1 + Ax,$$

where A is an arbitrary constant.

(b) The differential equation is $dy/dx = 2y/(x^2 + 1)$. In order to apply the separation of variables method, we need to exclude the case where $y = 0$. So for $y \neq 0$, on applying Procedure 2 we have

$$\int \frac{1}{y}\,dy = \int \frac{2}{x^2+1}\,dx.$$

Since $y \neq 0$, integration produces

$$\ln|y| = 2\arctan x + C,$$

where C is an arbitrary constant. On solving this equation for y, we obtain

$$y = \pm e^{2\arctan x + C} = \pm e^C e^{2\arctan x} = Be^{2\arctan x},$$

where $B = \pm e^C$ is a non-zero but otherwise arbitrary constant.

Examination of the differential equation shows that $B = 0$ also gives a solution (the zero function $y = 0$).

The general solution is therefore

$$y = Be^{2\arctan x},$$

where B is an arbitrary constant.

Solution to Exercise 9

(a) The given equation is $dP/dt = kP(1 - P/M)$. First, note that the constant functions $P = 0$ and $P = M$ are both solutions. Assuming that we are considering neither of these possibilities (we are certainly not interested in $P = 0$ since we know that $P_0 > 0$), we can use the separation of variables method to obtain

$$\frac{1}{k}\int \frac{1}{P(1-P/M)}\,dP = \int 1\,dt.$$

The integral on the left-hand side is of the form evaluated in Exercise 6(b), with $1/M$ in place of a. Hence we have

$$-\frac{1}{k}\ln\left|\frac{1}{P} - \frac{1}{M}\right| = t + C,$$

where C is an arbitrary constant. On solving for P, we find first that

$$\frac{1}{P} - \frac{1}{M} = \pm e^{-k(t+C)} = \pm e^{-kC}e^{-kt} = Be^{-kt},$$

where $B = \pm e^{-kC}$ is a non-zero but otherwise arbitrary constant. However, note that $B = 0$ corresponds to the constant solution $P = M$ already noted, so the restriction $B \neq 0$ may be dropped. Hence we obtain

$$P = \left(\frac{1}{M} + Be^{-kt}\right)^{-1} \qquad (e^{kt} \neq -MB),$$

where B is an arbitrary constant.

From the initial condition $P(0) = P_0$, we have

$$P_0 = \left(\frac{1}{M} + Be^{0}\right)^{-1}, \quad \text{so} \quad B = \frac{1}{P_0} - \frac{1}{M}.$$

The solution to the initial-value problem is therefore

$$P = \left(\frac{1}{M} + \left(\frac{1}{P_0} - \frac{1}{M}\right)e^{-kt}\right)^{-1},$$

which yields

$$P = \frac{M}{1 + (M/P_0 - 1)e^{-kt}}.$$

(b) As $t \to \infty$, we have $e^{-kt} \to 0$, and consequently the value of $P(t)$ approaches M. This is true whether the starting value P_0 is greater than or less than M.

Solution to Exercise 10

(a) The equation $dy/dx = x\sin x$ is linear, with $g(x) = 0$ (for all x) and $h(x) = x\sin x$.

(b) The equation $\dot{y} + y^2 = t$ is not linear (because of the y^2 term).

(c) The equation $x(dy/dx) + y = y^2$ is not linear (because of the y^2 term).

(d) The equation $(1+x^2)(dy/dx)+2xy=3x^2$ is linear, since we can divide through by $1+x^2$ to obtain

$$\frac{dy}{dx}+\frac{2xy}{1+x^2}=\frac{3x^2}{1+x^2},$$

which is of the defined form with $g(x)=2x/(1+x^2)$ and $h(x)=3x^2/(1+x^2)$.

Solution to Exercise 11

(a) The given equation is $dy/dx-y=e^x\sin x$. Comparison with equations (16) and (17) shows that the integrating factor is

$$p(x)=\exp\left(\int(-1)\,dx\right)=\exp(-x)=e^{-x}.$$

Multiplying through by $p(x)$ gives

$$e^{-x}\frac{dy}{dx}-e^{-x}y=\sin x.$$

Thus the differential equation can be rewritten as

$$\frac{d}{dx}(e^{-x}y)=\sin x.$$

On integrating, we find the general solution

$$e^{-x}y=-\cos x+C,$$

or equivalently,

$$y=e^x(C-\cos x),$$

where C is an arbitrary constant.

(b) The given equation, when rearranged into form (16), is $dy/dx-y/x=-1/x$. Thus $g(x)=-1/x$ and $h(x)=-1/x$, and the integrating factor is

$$p(x)=\exp\left(\int g(x)\,dx\right)=\exp\left(\int-\frac{1}{x}\,dx\right)$$
$$=\exp(-\ln x)=\exp\left(\ln\left(\frac{1}{x}\right)\right)=\frac{1}{x}.$$

Multiplying through by $p(x)$ gives

$$\frac{1}{x}\frac{dy}{dx}-\frac{1}{x^2}y=-\frac{1}{x^2}.$$

Thus the differential equation can be rewritten as

$$\frac{d}{dx}\left(\frac{y}{x}\right)=-\frac{1}{x^2}.$$

Hence

$$\frac{y}{x}=\int-\frac{1}{x^2}\,dx=\frac{1}{x}+C,$$

where C is an arbitrary constant.

After multiplying through by x, the general solution in explicit form is

$$y = 1 + Cx.$$

Solution to Exercise 12

(a) This requires the integrating factor method.

(b) This is best solved by direct integration.

(c) This can be solved by separation of variables or the integrating factor method.

(d) This requires the integrating factor method.

Solution to Exercise 13

For the initial-value problem

$$\frac{dy}{dx} = y, \quad y(0) = 1,$$

we have $x_0 = 0$, $Y_0 = y_0 = 1$ and $f(x_i, Y_i) = Y_i$. The step size is given as $h = 0.2$. Equation (25) with $i = 0$ gives

$$x_1 = x_0 + h = 0 + 0.2 = 0.2,$$

and equation (26) with $i = 0$ gives

$$Y_1 = Y_0 + h\, f(x_0, Y_0) = 1 + 0.2 \times 1 = 1.2.$$

Applying equations (25) and (26) in turn for $i = 1, 2, 3, 4$, we obtain the following table.

i	x_i	Y_i	$f(x_i, Y_i) = Y_i$	$Y_{i+1} = Y_i + h\, f(x_i, Y_i)$
0	0	1	1	1.2
1	0.2	1.2	1.2	1.44
2	0.4	1.44	1.44	1.728
3	0.6	1.728	1.728	2.073 6
4	0.8	2.073 6	2.073 6	2.488 32
5	1.0	2.488 32		

The approximation to $y(1)$ is 2.488 32.

Solution to Exercise 14

Since we are told that for a sufficiently small step size h, the absolute error is proportional to h, we can deduce from the last row of Table 2 that there exists a constant k such that

$$0.000\,136 = 0.0001k,$$

so $k = 1.36$. In order to determine $y(1)$ correct to five decimal places, h must be such that

$$1.36h < 5 \times 10^{-6}$$

or

$$h < \frac{5 \times 10^{-6}}{1.36} \simeq 3.7 \times 10^{-6}.$$

So a suitable choice of h would be $10^{-6} = 0.000\,001$.

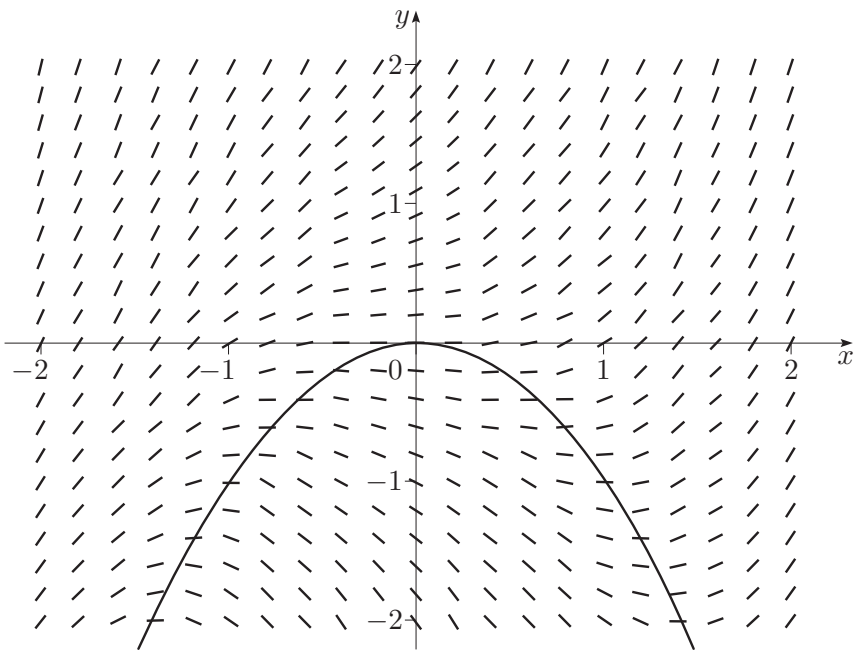
(In fact, using this value of h gives an approximation to $y(1)$ of 2.718 280, which *is* correct to five decimal places.)

Solution to Exercise 15

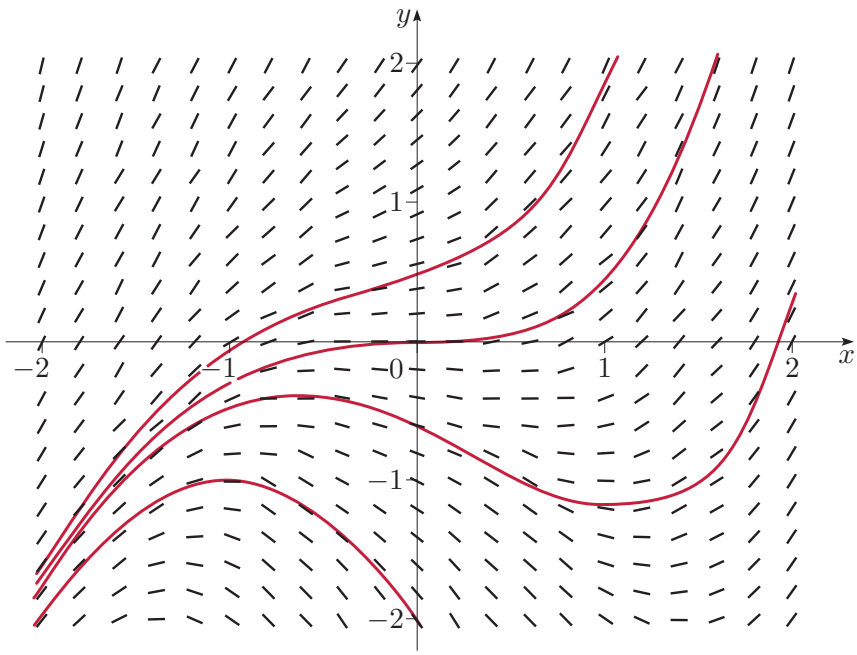
(a) The slope defined by the direction field $f(x, y) = y + x^2$ is zero when $y = -x^2$, which is a parabola in the lower half-plane with vertex at the origin. Below this parabola we have $y < -x^2$ and $f(x, y) < 0$, while above the parabola we have $y > -x^2$ and $f(x, y) > 0$. Thus all slopes for points of the plane below the parabola $y = -x^2$ are negative, and all slopes for points above it are positive.

Also, if x is fixed, then $f(x, y) = y + x^2$ is an increasing function as y increases. If instead y is fixed, then for $x > 0$, $f(x, y)$ increases as x increases, and for $x < 0$, $f(x, y)$ increases as x becomes more negative. These observations indicate that the slope given by the direction field increases as we move from bottom to top along any vertical line, whereas on moving along any horizontal line, the slope increases with distance from the y-axis.

(b) The features described in the solution to part (a) are all apparent on the direction field diagram. This direction field diagram is repeated below, with the parabola $y = -x^2$ superimposed on it. (Note that this parabola does not represent a solution of the differential equation.)



(c) It appears from the direction field that there are several types of solution. Any solution whose graph cuts the y-axis above the origin has positive slope at all points. The solution graph that passes through the origin has zero slope there, but positive slope everywhere else. Any solution graph that cuts the y-axis below the origin has a maximum (where it meets $y = -x^2$ for $x < 0$). Some of these graphs also have a minimum (where they meet $y = -x^2$ for $x > 0$). Others have no minimum (though this is not clear from the diagram given). A solution graph of each type is sketched below.



(d) The initial-value problem is

$$\frac{dy}{dx} = y + x^2, \quad y(-1) = -0.2.$$

From equations (25) and (26), the necessary formulas are

$$x_{i+1} = x_i + h,$$
$$Y_{i+1} = Y_i + h\, f(x_i, Y_i).$$

For the current problem, $x_0 = -1$, $Y_0 = y_0 = -0.2$, $f(x_i, Y_i) = Y_i + x_i^2$ and $h = 0.1$. The particular formulas needed here are therefore

$$x_{i+1} = x_i + 0.1, \quad \text{where } x_0 = -1,$$
$$Y_{i+1} = Y_i + 0.1(Y_i + x_i^2), \quad \text{where } Y_0 = -0.2.$$

The second of these formulas can also be written as

$$Y_{i+1} = 1.1Y_i + 0.1x_i^2, \quad \text{where } Y_0 = -0.2.$$

Solution to Exercise 16

(a) Equations (i), (ii), (iii), (vi), (vii) and (viii) are linear and constant-coefficient. (Equation (v) is non-linear; (iv) is linear but not constant-coefficient.)

(b) Of the linear constant-coefficient equations, only (iii) and (viii) are homogeneous.

(c) In equations (i)–(v) the (dependent, independent) variable pairs are all (y, x). In equations (vi), (vii) and (viii) they are (t, θ), (x, t) and (x, t), respectively.

Solution to Exercise 17

(a) $\lambda^2 - 5\lambda + 6 = 0$

(b) $\lambda^2 - 9 = 0$

(c) $\lambda^2 + 2\lambda = 0$

Solution to Exercise 18

(a) The auxiliary equation is $\lambda^2 + 5\lambda + 6 = 0$. Solving this by factorisation as $(\lambda + 2)(\lambda + 3) = 0$ gives the roots $\lambda_1 = -2$ and $\lambda_2 = -3$. The general solution is therefore

$$y = Ce^{-2x} + De^{-3x},$$

where C and D are arbitrary constants.

(b) The auxiliary equation is $2\lambda^2 + 3\lambda = 0$. This can be factorised as $\lambda(2\lambda + 3) = 0$, so its roots are $\lambda_1 = 0$ and $\lambda_2 = -\frac{3}{2}$. The general solution is therefore

$$y = Ce^0 + De^{-3x/2} = C + De^{-3x/2},$$

where C and D are arbitrary constants.

(c) The auxiliary equation is $\lambda^2 - 4 = 0$, that is, $\lambda^2 = 4$, so its roots are $\lambda_1 = -2$ and $\lambda_2 = 2$. The general solution is therefore

$$z = Ce^{-2u} + De^{2u},$$

where C and D are arbitrary constants.

Solution to Exercise 19

(a) The auxiliary equation is $\lambda^2 + 2\lambda + 1 = 0$, which can be factorised as $(\lambda + 1)^2 = 0$, giving equal roots $\lambda_1 = \lambda_2 = -1$. The general solution is therefore

$$y = (C + Dx)e^{-x},$$

where C and D are arbitrary constants.

(b) The auxiliary equation $\lambda^2 - 4\lambda + 4 = 0$ factorises as $(\lambda - 2)^2 = 0$, which has equal roots $\lambda_1 = \lambda_2 = 2$. The general solution is therefore

$$s = (C + Dt)e^{2t},$$

where C and D are arbitrary constants.

Solution to Exercise 20

(a) The auxiliary equation is $\lambda^2 + 4\lambda + 8 = 0$, which has solutions

$$\lambda = \frac{-4 \pm \sqrt{16 - 32}}{2} = -2 \pm 2i.$$

The general solution is therefore

$$y = e^{-2x}(C\cos 2x + D\sin 2x),$$

where C and D are arbitrary constants.

(b) The auxiliary equation is $\lambda^2 + 9 = 0$, which has solutions

$$\lambda = \pm 3i.$$

The general solution is therefore

$$\theta = e^{0}(C\cos 3t + D\sin 3t) = C\cos 3t + D\sin 3t,$$

where C and D are arbitrary constants.

Solution to Exercise 21

(a) The auxiliary equation is $\lambda^2 + 4 = 0$, which has solutions $\lambda = \pm 2i$. The general solution is therefore

$$y = C\cos 2x + D\sin 2x,$$

where C and D are arbitrary constants.

(b) The auxiliary equation is $\lambda^2 - 6\lambda + 8 = 0$, which has solutions $\lambda_1 = 4$ and $\lambda_2 = 2$. The general solution is therefore

$$u = Ce^{4x} + De^{2x},$$

where C and D are arbitrary constants.

(c) The auxiliary equation is $\lambda^2 + 2\lambda = 0$, which has solutions $\lambda_1 = 0$ and $\lambda_2 = -2$. The general solution is therefore

$$y = C + De^{-2x},$$

where C and D are arbitrary constants.

(d) The auxiliary equation is $\lambda^2 - 2\lambda + 1 = 0$, which has solutions $\lambda_1 = \lambda_2 = 1$. The general solution is therefore

$$y = (C + Dx)e^{x},$$

where C and D are arbitrary constants.

(e) The auxiliary equation is $\lambda^2 - \omega^2 = 0$, which has solutions $\lambda = \pm\omega$. The general solution is therefore

$$y = Ce^{\omega x} + De^{-\omega x},$$

where C and D are arbitrary constants.

(f) The auxiliary equation is $\lambda^2 + 4\lambda + 29 = 0$, which has solutions $\lambda = -2 \pm 5i$. The general solution is therefore

$$e^{-2x}(C\cos 5x + D\sin 5x),$$

where C and D are arbitrary constants.

Solution to Exercise 22

The auxiliary equation is $\lambda^2 + g/l = 0$, which has solutions $\lambda = \pm i\sqrt{g/l}$. The general solution is therefore

$$\theta = C\cos\left(\sqrt{\frac{g}{l}}\,t\right) + D\sin\left(\sqrt{\frac{g}{l}}\,t\right).$$

This problem will be discussed again in Unit 10.

Solution to Exercise 23

We could check this directly, by substituting $y = y_{\mathrm{p}_1} - y_{\mathrm{p}_2}$ into the associated homogeneous equation. However, it is easier to appeal to the principle of superposition. Since y_{p_1} and y_{p_2} both satisfy

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x),$$

Theorem 2 shows that the combination $y = y_{\mathrm{p}_1} - y_{\mathrm{p}_2}$ indeed satisfies

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = f(x) - f(x) = 0,$$

as required.

Solution to Exercise 24

(a) The associated homogeneous equation is $d^2y/dx^2 + 4y = 0$. The complementary function (see Exercise 21(a)) is $y_{\mathrm{c}} = C\cos 2x + D\sin 2x$, where C and D are arbitrary constants.

Trying a solution of the form $y_{\mathrm{p}} = p$, where p is a constant, in the original equation $d^2y/dx^2 + 4y = 8$ gives $0 + 4p = 8$, so $p = 2$. Thus a particular integral is $y_{\mathrm{p}} = 2$.

By Procedure 6, the general solution is

$$y = C\cos 2x + D\sin 2x + 2.$$

(b) The associated homogeneous equation is $d^2y/dx^2 - 3\,dy/dx + 2y = 0$. The complementary function (see Example 8) is $y_{\mathrm{c}} = Ce^x + De^{2x}$, where C and D are arbitrary constants.

Trying a solution of the form $y_{\mathrm{p}} = p$ in the original equation $d^2y/dx^2 - 3\,dy/dx + 2y = 6$ gives $0 - 0 + 2p = 6$, so $p = 3$. Thus a particular integral is $y_{\mathrm{p}} = 3$.

By Procedure 6, the general solution is

$$y = Ce^x + De^{2x} + 3.$$

Solution to Exercise 25

(a) Substituting $y = p_1x + p_0$ and its derivatives into the differential equation gives

$$0 - 2p_1 + 2(p_1x + p_0) = 2p_1x + (2p_0 - 2p_1) = 2x + 3.$$

Equating coefficients gives $p_1 = 1$, $p_0 = \frac{5}{2}$. Therefore a particular integral is

$$y_\mathrm{p} = x + \tfrac{5}{2}.$$

(b) Substituting $y = p_1x + p_0$ and its derivatives into the differential equation gives

$$0 + 2p_1 + (p_1x + p_0) = p_1x + (2p_1 + p_0) = 2x.$$

Hence $p_1 = 2$, $p_0 = -4$, and a particular integral is

$$y_\mathrm{p} = 2x - 4.$$

Solution to Exercise 26

We try $y = p_2t^2 + p_1t + p_0$, which has derivatives $\dot{y} = 2p_2t + p_1$, $\ddot{y} = 2p_2$. Substituting these into the differential equation gives

$$2p_2 - (p_2t^2 + p_1t + p_0) = -p_2t^2 - p_1t + (2p_2 - p_0) = t^2.$$

Hence $p_2 = -1$, $p_1 = 0$, $p_0 = -2$, and a particular integral is

$$y_\mathrm{p} = -t^2 - 2.$$

Solution to Exercise 27

We try a solution of the form $y = pe^{-x}$, which has derivatives $dy/dx = -pe^{-x}$, $d^2y/dx^2 = pe^{-x}$. Substituting these into the differential equation gives

$$2pe^{-x} + 2pe^{-x} + pe^{-x} = 5pe^{-x} = 2e^{-x}.$$

Hence $p = \frac{2}{5}$, and a particular integral is

$$y_\mathrm{p} = \tfrac{2}{5}e^{-x}.$$

Solution to Exercise 28

We try $y = p\cos 3t + q\sin 3t$, which has derivatives

$$\frac{dy}{dt} = -3p\sin 3t + 3q\cos 3t, \quad \frac{d^2y}{dt^2} = -9p\cos 3t - 9q\sin 3t.$$

Substituting into the differential equation gives

$$\begin{aligned}&(-9p\cos 3t - 9q\sin 3t) - (-3p\sin 3t + 3q\cos 3t)\\ &\quad = -(9p + 3q)\cos 3t + (3p - 9q)\sin 3t\\ &\quad = \cos 3t + \sin 3t.\end{aligned}$$

Hence we have a pair of simultaneous equations

$$\begin{aligned}-9p - 3q &= 1,\\ 3p - 9q &= 1.\end{aligned}$$

Adding three times the second equation to the first, to eliminate p, gives $-30q = 4$, so $q = -\frac{2}{15}$, whence $p = -\frac{1}{15}$. A particular integral is thus

$$y_{\mathrm{p}} = -\tfrac{1}{15}\cos 3t - \tfrac{2}{15}\sin 3t.$$

Solution to Exercise 29

(a) Try $y = pe^{2x}$.

(b) Try $y = p\cos 4x + q\sin 4x$.

Solution to Exercise 30

(a) The complementary function is $y_{\mathrm{c}} = e^{-x}(C\cos x + D\sin x)$, where C and D are arbitrary constants.

To find a particular integral, try $y = p_0$. Substituting into the differential equation gives

$$0 + 0 + 2p_0 = 2p_0 = 4.$$

Hence $p_0 = 2$, and a particular integral is $y_{\mathrm{p}} = 2$.

Therefore the general solution is

$$y = e^{-x}(C\cos x + D\sin x) + 2.$$

(b) The complementary function is $\theta_{\mathrm{c}} = C + De^{-3t}$, where C and D are arbitrary constants.

To find a particular integral, try $\theta = p\cos 2t + q\sin 2t$. Differentiating gives

$$\frac{d\theta}{dt} = -2p\sin 2t + 2q\cos 2t, \quad \frac{d^2\theta}{dt^2} = -4p\cos 2t - 4q\sin 2t.$$

Substituting into the differential equation gives

$$\begin{aligned}&(-4p\cos 2t - 4q\sin 2t) + 3(-2p\sin 2t + 2q\cos 2t)\\ &\quad = (6q - 4p)\cos 2t - (4q + 6p)\sin 2t\\ &\quad = 13\cos 2t.\end{aligned}$$

Comparing the coefficients of $\cos 2t$ and $\sin 2t$ gives a pair of simultaneous equations to solve:

$$\begin{aligned}-4p + 6q &= 13,\\ -6p - 4q &= 0.\end{aligned}$$

Hence $p = -1$, $q = \frac{3}{2}$, and a particular integral is $\theta_{\mathrm{p}} = -\cos 2t + \frac{3}{2}\sin 2t$.

Therefore the general solution is

$$\theta = C + De^{-3t} - \cos 2t + \tfrac{3}{2}\sin 2t.$$

Solution to Exercise 31

Putting the equation into the usual form and using $R = S = Q = 1$ gives

$$y'' - y = -\tfrac{1}{2}(l - x)x = -\tfrac{1}{2}lx + \tfrac{1}{2}x^2.$$

The associated homogeneous equation is $y'' - y = 0$, which has auxiliary equation $\lambda^2 - 1 = 0$. This has roots $\lambda = \pm 1$, so the complementary function is

$$y_c = Ce^x + De^{-x},$$

where C and D are arbitrary constants.

To obtain a particular integral, we try a function of the form $y = p_2x^2 + p_1x + p_0$. Its derivatives are $y' = 2p_2x + p_1$, $y'' = 2p_2$. Substituting into the differential equation gives

$$\begin{aligned} 2p_2 - (p_2x^2 + p_1x + p_0) &= -p_2x^2 - p_1x + (2p_2 - p_0) \\ &= \tfrac{1}{2}x^2 - \tfrac{1}{2}lx. \end{aligned}$$

Hence $p_2 = -\frac{1}{2}$, $p_1 = \frac{1}{2}l$, $p_0 = -1$, and a particular integral is

$$y_p = -\tfrac{1}{2}x^2 + \tfrac{1}{2}lx - 1.$$

Therefore the general solution is

$$y = Ce^x + De^{-x} - \tfrac{1}{2}x^2 + \tfrac{1}{2}lx - 1.$$

Solution to Exercise 32

(a) From Example 8, the associated homogeneous equation has general solution $y = Ce^x + De^{2x}$, where C and D are arbitrary constants, and the trial solution $y = pe^x$ suggested by Procedure 7 is a solution of this equation (with $C = p$, $D = 0$). So we try $y = pxe^x$ instead. Differentiating twice gives

$$\frac{dy}{dx} = pe^x + pxe^x = p(1 + x)e^x,$$
$$\frac{d^2y}{dx^2} = pe^x + p(1 + x)e^x = p(2 + x)e^x.$$

Substituting into the differential equation gives

$$p(2 + x)e^x - 3p(1 + x)e^x + 2pxe^x = -pe^x = 4e^x.$$

Hence $p = -4$, and a particular integral is

$$y_p = -4xe^x.$$

(b) From Exercise 18(b), the associated homogeneous equation has general solution $y = C + De^{-3x/2}$, where C and D are arbitrary constants, and the trial solution $y = p_0$ suggested by Procedure 7 is a solution of this equation (with $C = p_0$, $D = 0$). So we try $y = p_0x$. Differentiating twice gives

$$\frac{dy}{dx} = p_0, \quad \frac{d^2y}{dx^2} = 0.$$

Substituting into the differential equation gives $3p_0 = 1$, so $p_0 = \frac{1}{3}$, and a particular integral is

$$y_\mathrm{p} = \tfrac{1}{3}x.$$

Solution to Exercise 33

The auxiliary equation of the differential equation $m\ddot{x} + r\dot{x} = mg$ is

$$m\lambda^2 + r\lambda = 0,$$

with solutions $\lambda = 0$ and $\lambda = -r/m$. The complementary function is therefore

$$x_\mathrm{c} = C + De^{-rt/m},$$

where C and D are arbitrary constants.

The inhomogeneous term is mg, so Procedure 7 suggests a trial solution $x = p_0$. However, this is a solution of the associated homogeneous equation (with $C = p_0$, $D = 0$). Hence we try $x = p_0 t$ instead, where t is the independent variable in this problem. Differentiating and substituting gives

$$rp_0 = mg,$$

so

$$p_0 = \frac{mg}{r}.$$

Hence a particular integral is

$$x_\mathrm{p} = \frac{mgt}{r},$$

and the general solution is

$$x = C + De^{-rt/m} + \frac{mgt}{r}.$$

Solution to Exercise 34

From Exercise 21(d), the associated homogeneous equation has general solution $y = (C + Dx)e^x$, where C and D are arbitrary constants. So not only is the trial solution $y = pe^x$ suggested by Procedure 7 a solution of the associated homogeneous differential equation (with $C = p$, $D = 0$), but so is $y = pxe^x$ (with $C = 0$, $D = p$). So we try $y = px^2e^x$. Differentiating twice gives

$$\frac{dy}{dx} = 2pxe^x + px^2e^x = p(2x + x^2)e^x,$$
$$\frac{d^2y}{dx^2} = p(2 + 2x)e^x + p(2x + x^2)e^x = p(2 + 4x + x^2)e^x.$$

Substituting into the differential equation gives

$$p(2 + 4x + x^2)e^x - 2p(2x + x^2)e^x + px^2e^x = 2pe^x = e^x.$$

Hence $p = \frac{1}{2}$, and a particular integral is

$$y_\mathrm{p} = \tfrac{1}{2}x^2e^x.$$

Solution to Exercise 35

(a) From Exercise 34, $y_p = \frac{1}{2}x^2e^x$ is a particular integral for

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + y = e^x.$$

So, using the principle of superposition, we can find a particular integral for the given differential equation if we can find one for

$$\frac{d^2y}{dx} - 2\frac{dy}{dx} + y = -3e^{2x}.$$

We try $y = pe^{2x}$, which has derivatives

$$\frac{dy}{dx} = 2pe^{2x}, \quad \frac{d^2y}{dx^2} = 4pe^{2x}.$$

Substituting into the differential equation gives

$$4pe^{2x} - 4pe^{2x} + pe^{2x} = pe^{2x} = -3e^{2x}.$$

Hence $p = -3$, and $y_p = -3e^{2x}$ is a particular integral for the differential equation with right-hand side $-3e^{2x}$.

Thus, using the principle of superposition, a particular integral for the given differential equation is

$$y_p = 4\left(\tfrac{1}{2}x^2e^x\right) - 3e^{2x} = 2x^2e^x - 3e^{2x}.$$

(b) This time we do not have a particular integral for any part of the right-hand-side function, so we need to start from scratch.

First consider the $12\cos 2t$ term on the right-hand side, and try $x = p\cos 2t + q\sin 2t$ as a trial solution. This has derivatives

$$\frac{dx}{dt} = -2p\sin 2t + 2q\cos 2t, \quad \frac{d^2x}{dt^2} = -4p\cos 2t - 4q\sin 2t.$$

Substituting into the differential equation gives

$$\begin{aligned} &2(-4p\cos 2t - 4q\sin 2t) + 3(-2p\sin 2t + 2q\cos 2t) \\ &\qquad + 2(p\cos 2t + q\sin 2t) \\ &= 6(q-p)\cos 2t - 6(p+q)\sin 2t \\ &= 12\cos 2t. \end{aligned}$$

So $p + q = 0$, $q - p = 2$, hence $p = -1$, $q = 1$, and a particular integral is

$$x_p = -\cos 2t + \sin 2t.$$

Now consider the 10 term, and try $x = p_0$. Substituting into the differential equation gives $2p_0 = 10$, so $p_0 = 5$, and a particular integral is

$$x_p = 5.$$

Therefore, using the principle of superposition, a particular integral for the differential equation with $f(t) = 12\cos 2t + 10$ is

$$x_p = -\cos 2t + \sin 2t + 5.$$

Solution to Exercise 36

(a) From Exercise 21(a), the complementary function is

$$\theta_{\mathrm{c}} = C\cos 2t + D\sin 2t,$$

where C and D are arbitrary constants.

To find a particular integral, try $\theta = p_1 t + p_0$. Substituting this and its derivatives into the differential equation gives

$$4(p_1 t + p_0) = 2t.$$

Hence $p_1 = \frac{1}{2}$, $p_0 = 0$, and a particular integral is

$$\theta_{\mathrm{p}} = \tfrac{1}{2}t.$$

Therefore the general solution is

$$\theta = C\cos 2t + D\sin 2t + \tfrac{1}{2}t.$$

(b) The auxiliary equation is $\lambda^2 + 4\lambda + 5 = 0$, which has solutions $\lambda = -2 \pm i$. So (using Procedure 5) the complementary function is

$$u_{\mathrm{c}} = e^{-2t}(C\cos t + D\sin t),$$

where C and D are arbitrary constants.

To find a particular integral, try $u = p_0$. Substituting gives $5p_0 = 5$. Hence $p_0 = 1$, and a particular integral is

$$u_{\mathrm{p}} = 1.$$

Therefore the general solution is

$$u = e^{-2t}(C\cos t + D\sin t) + 1.$$

(c) The auxiliary equation is $3\lambda^2 - 2\lambda - 1 = 0$, which has solutions $\lambda_1 = 1$ and $\lambda_2 = -\frac{1}{3}$. So the complementary function is

$$Y_{\mathrm{c}} = Ce^{x} + De^{-x/3},$$

where C and D are arbitrary constants.

Consider first the e^{2x} term on the right-hand side of the equation. To find a particular integral, try $Y = pe^{2x}$. The derivatives are $dY/dx = 2pe^{2x}$ and $d^2Y/dx^2 = 4pe^{2x}$. Substituting gives

$$3(4pe^{2x}) - 2(2pe^{2x}) - pe^{2x} = 7pe^{2x} = e^{2x}.$$

Hence $p = \frac{1}{7}$, and a particular integral is

$$Y_{\mathrm{p}} = \tfrac{1}{7}e^{2x}.$$

Now consider the 3 term on the right-hand side of the equation, and try $Y = p_0$. Substituting gives $-p_0 = 3$, so $p_0 = -3$, and a particular integral is

$$Y_{\mathrm{p}} = -3.$$

Therefore, using the principle of superposition, a particular integral for the differential equation with $f(x) = e^{2x} + 3$ is

$$Y_{\mathrm{p}} = \tfrac{1}{7}e^{2x} - 3.$$

Therefore the general solution is

$$Y = Ce^{x} + De^{-x/3} + \tfrac{1}{7}e^{2x} - 3.$$

(d) From Exercise 18(c), the complementary function is

$$y_{\mathrm{c}} = Ce^{-2x} + De^{2x},$$

where C and D are arbitrary constants.

To find a particular integral, since e^{-2x} is a solution of the associated homogeneous equation, try $y = pxe^{-2x}$. The derivatives are $dy/dx = p(1-2x)e^{-2x}$ and $d^2y/dx^2 = 4p(x-1)e^{-2x}$. Substituting gives

$$4p(x-1)e^{-2x} - 4pxe^{-2x} = -4pe^{-2x} = e^{-2x}.$$

Hence $p = -\frac{1}{4}$, and a particular integral is

$$y_{\mathrm{p}} = -\tfrac{1}{4}xe^{-2x}.$$

Therefore the general solution is

$$y = Ce^{-2x} + De^{2x} - \tfrac{1}{4}xe^{-2x}.$$

(e) From Exercise 21(a), the complementary function is

$$y_{\mathrm{c}} = C\cos 2x + D\sin 2x,$$

where C and D are arbitrary constants.

To find a particular integral, we note that, from part (a) of this exercise, a particular integral for $d^2y/dx^2 + 4y = 2x$ is $y_{\mathrm{p}} = \frac{1}{2}x$, so for a right-hand side of $3x$ we will have $y_{\mathrm{p}} = \frac{3}{2}(\frac{1}{2}x) = \frac{3}{4}x$. So we need to consider only the $\sin 2x$ term, and then use the principle of superposition.

For this term, noting the form of the complementary function, try $y = x(p\cos 2x + q\sin 2x)$. The derivatives are

$$\frac{dy}{dx} = (p + 2qx)\cos 2x + (q - 2px)\sin 2x,$$
$$\frac{d^2y}{dx^2} = (4q - 4px)\cos 2x - (4p + 4qx)\sin 2x.$$

Substituting gives

$$\begin{aligned}&(4q - 4px)\cos 2x - (4p + 4qx)\sin 2x + 4x(p\cos 2x + q\sin 2x)\\ &\quad = 4q\cos 2x - 4p\sin 2x\\ &\quad = \sin 2x.\end{aligned}$$

Hence $p = -\frac{1}{4}$, $q = 0$, and a particular integral is

$$y_{\mathrm{p}} = -\tfrac{1}{4}x\cos 2x.$$

Therefore, using the principle of superposition, a particular integral for the given differential equation is

$$y_{\mathrm{p}} = \tfrac{3}{4}x - \tfrac{1}{4}x\cos 2x.$$

Thus the general solution is

$$y = C\cos 2x + D\sin 2x + \tfrac{3}{4}x - \tfrac{1}{4}x\cos 2x.$$

(f) From Example 8, the complementary function is

$$y_c = Ce^x + De^{2x},$$

where C and D are arbitrary constants.

Consider first the $2e^x$ term on the right-hand side of the equation. To find a particular integral, since e^x appears in the complementary function, try $y = pxe^x$, which has derivatives

$$\frac{dy}{dx} = p(1+x)e^x, \quad \frac{d^2y}{dx^2} = p(2+x)e^x.$$

Substituting into the differential equation gives

$$p(2+x)e^x - 3p(1+x)e^x + 2pxe^x = -pe^x = 2e^x.$$

Hence $p = -2$, and a particular integral is

$$y_p = -2xe^x.$$

Now consider the $-5e^{2x}$ term on the right-hand side of the equation. To find a particular integral, since e^{2x} appears in the complementary function, try $y = pxe^{2x}$, which has derivatives

$$\frac{dy}{dx} = p(1+2x)e^{2x}, \quad \frac{d^2y}{dx^2} = p(4+4x)e^{2x}.$$

Substituting into the differential equation gives

$$p(4+4x)e^{2x} - 3p(1+2x)e^{2x} + 2pxe^{2x} = pe^{2x} = -5e^{2x}.$$

Hence $p = -5$, and a particular integral is

$$y_p = -5xe^{2x}.$$

Therefore, using the principle of superposition, a particular integral for the differential equation with $f(x) = 2e^x - 5e^{2x}$ is

$$y_p = -2xe^x - 5xe^{2x}.$$

Therefore the general solution is

$$y = Ce^x + De^{2x} - 2xe^x - 5xe^{2x}.$$

Solution to Exercise 37

(a) From Exercise 20(b), the general solution is

$$u = C\cos 3t + D\sin 3t,$$

where C and D are arbitrary constants. Its derivative is

$$u' = -3C\sin 3t + 3D\cos 3t.$$

Substituting the initial condition $t = \frac{\pi}{2}$, $u = 0$ into the general solution gives $D = 0$. Substituting the initial condition $t = \frac{\pi}{2}$, $u' = 1$ into the derivative gives $C = \frac{1}{3}$. Hence the required particular solution is

$$u = \tfrac{1}{3}\cos 3t.$$

(b) From Exercise 32(a), the general solution is

$$y = Ce^x + De^{2x} - 4xe^x,$$

where C and D are arbitrary constants. Its derivative is

$$y' = Ce^x + 2De^{2x} - 4(1 + x)e^x.$$

Substituting the initial condition $x = 0$, $y = 4$ into the general solution gives $C + D = 4$. Substituting the initial condition $x = 0$, $y' = 2$ into the derivative gives $C + 2D - 4 = 2$. Hence $C = 2$, $D = 2$, and the required particular solution is

$$y = 2e^x + 2e^{2x} - 4xe^x = (2 + 4x)e^x + 2e^{2x}.$$

(c) From Exercises 21(d) and 35(a), the general solution is

$$y = (C + Dx)e^x + 2x^2e^x - 3e^{2x},$$

where C and D are arbitrary constants. Its derivative is

$$y' = (C + D + Dx)e^x + (4x + 2x^2)e^x - 6e^{2x}.$$

Substituting the initial condition $x = 0$, $y = 4$ into the general solution gives $C - 3 = 4$. Substituting the initial condition $x = 0$, $y' = -1$ into the derivative gives $C + D - 6 = -1$. Hence $C = 7$, $D = -2$, and the required particular solution is

$$y = (7 - 2x)e^x + 2x^2e^x - 3e^{2x} = (7 - 2x + 2x^2)e^x - 3e^{2x}.$$

Solution to Exercise 38

From Exercise 32(a), the general solution is

$$y = Ce^x + De^{2x} - 4xe^x,$$

where C and D are arbitrary constants, and its derivative is

$$y' = Ce^x + 2De^{2x} - 4(1 + x)e^x.$$

Substituting the boundary condition $x = 0$, $y' = 2$ into the derivative gives $C + 2D - 4 = 2$. Substituting $x = 1$, $y = 0$ into the general solution gives $Ce + De^2 - 4e = 0$, which can be rearranged to give $C + eD = 4$. Hence $C = (8 - 6e)/(2 - e)$, $D = 2/(2 - e)$, and the required particular solution is

$$y = \frac{8 - 6e}{2 - e}e^x + \frac{2}{2 - e}e^{2x} - 4xe^x.$$

Solution to Exercise 39

From Exercise 31, the general solution of the differential equation is $y = Ce^x + De^{-x} - \frac{1}{2}x^2 + \frac{1}{2}lx - 1$, which for $l = 2$ becomes

$$y = Ce^x + De^{-x} - \tfrac{1}{2}x^2 + x - 1,$$

where C and D are arbitrary constants.

The boundary conditions, resulting from the beam resting on supports at its two ends, are $y(0) = 0$, $y(2) = 0$.

Substitution of these into the general solution gives $C + D - 1 = 0$ and $Ce^2 + De^{-2} - 1 = 0$. Multiplying the second equation by e^2 gives $C + D = 1$ and $Ce^4 + D = e^2$ as the equations to solve. Subtracting the equations gives $C(e^4 - 1) = e^2 - 1$, thus

$$C = \frac{e^2 - 1}{e^4 - 1} = \frac{e^2 - 1}{(e^2 + 1)(e^2 - 1)} = \frac{1}{e^2 + 1},$$

$$D = 1 - C = 1 - \frac{1}{e^2 + 1} = \frac{e^2 + 1 - 1}{e^2 + 1} = \frac{e^2}{e^2 + 1}.$$

Hence the required particular solution is

$$y = \frac{1}{e^2 + 1}(e^x + e^{2-x}) - \tfrac{1}{2}x^2 + x - 1.$$

You can check that this is symmetric about the centre of the beam by replacing x with $2 - x$.

At the centre of the beam, $x = 1$, so $y \simeq 0.148$. The displacement or 'sag' at the centre of the beam is approximately 0.148 m or about 14.8 cm.

Solution to Exercise 40

(a) The differential equation is the same in each case, and from Exercise 21(a) its general solution is

$$u = C\cos 2x + D\sin 2x,$$

where C and D are arbitrary constants, and the derivative is

$$u' = -2C\sin 2x + 2D\cos 2x.$$

In this case, we have an initial-value problem.

The condition $u(0) = 1$ gives $C = 1$. The condition $u'(0) = 0$ gives $D = 0$. The required solution is therefore

$$u = \cos 2x.$$

(b) This is a boundary-value problem.

The condition $u(0) = 0$ gives $C = 0$. The condition $u\left(\frac{\pi}{2}\right) = 0$ gives $C = 0$ also. D therefore remains arbitrary, so there is an infinite number of solutions, of the form

$$u = D\sin 2x.$$

(c) This is an initial-value problem.

The condition $u\left(\frac{\pi}{2}\right) = 0$ gives $C = 0$. The condition $u'\left(\frac{\pi}{2}\right) = 0$ gives $D = 0$. The required solution is therefore the zero function

$$u = 0.$$

(Alternatively, since the differential equation is homogeneous and the initial values are both equal to zero, by the remarks after Theorem 4, the solution is the zero function $u = 0$.)

(d) This is a boundary-value problem.

The condition $u(-\pi) = 1$ gives $C = 1$. The condition $u\left(\frac{\pi}{4}\right) = 2$ gives $D = 2$. The required solution is therefore

$$u = \cos 2x + 2\sin 2x.$$

(e) This is a boundary-value problem.

The condition $u'(0) = 0$ gives $D = 0$. The condition $u'\left(\frac{\pi}{4}\right) = 1$ gives $C = -\frac{1}{2}$. The required solution is therefore

$$u = -\tfrac{1}{2}\cos 2x.$$

Solution to Exercise 41

(a) This is an initial-value problem, therefore by Theorem 4 it has a unique solution.

The general solution is

$$u = e^{-2t}(C\cos t + D\sin t),$$

where C and D are arbitrary constants. Its derivative is

$$u' = e^{-2t}\big((-2C + D)\cos t - (C + 2D)\sin t\big).$$

The condition $u(0) = 0$ gives $C = 0$. The condition $u'(0) = 2$ gives $D = 2$. The solution is therefore

$$u = 2e^{-2t}\sin t.$$

(b) This is an initial-value problem, therefore by Theorem 4 it has a unique solution.

The differential equation is homogeneous and the initial values are both equal to zero. Hence the solution is the zero function $y = 0$.

(c) This is a boundary-value problem, which may have no solution, a unique solution, or an infinite number of solutions.

The complementary function is

$$x_{\mathrm{c}} = C\cos 3t + D\sin 3t,$$

where C and D are arbitrary constants.

To find a particular integral, try $x = p_1 t + p_0$. Substituting into the differential equation gives

$$9(p_1 t + p_0) = 3(1 - \pi t).$$

Hence $p_1 = -\frac{\pi}{3}$, $p_0 = \frac{1}{3}$, and a particular integral is

$$x_{\mathrm{p}} = -\tfrac{\pi}{3}t + \tfrac{1}{3}.$$

Therefore the general solution is

$$x = C\cos 3t + D\sin 3t - \tfrac{\pi}{3}t + \tfrac{1}{3},$$

and its derivative is

$$\dot{x} = -3C\sin 3t + 3D\cos 3t - \tfrac{\pi}{3}.$$

The condition $x(0) = \frac{1}{3}$ gives $C = 0$. The condition $\dot{x}\left(\frac{\pi}{3}\right) = 0$ gives $D = -\frac{\pi}{9}$. The solution is therefore

$$x = -\tfrac{\pi}{9}\sin 3t - \tfrac{\pi}{3}t + \tfrac{1}{3}.$$

Acknowledgements

Grateful acknowledgement is made to the following source:

Figure 10: Uploaded by ‘Soerfm’, from
http://upload.wikimedia.org/wikipedia/commons/6/60/Leonhard_Euler_2.jpg.

Every effort has been made to contact copyright holders. If any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

Unit 2

Vector algebra and statics

Introduction

This module discusses *Newtonian mechanics*. Isaac Newton (Figure 1) was the great English mathematician whose name is given to this subject. His *Philosophiae Naturalis Principia Mathematica* of 1687 (*Mathematical Principles of Natural Philosophy*, known as the *Principia* for short) incorporates one of the most celebrated examples of mathematical modelling. It was in the *Principia* that Newton laid down the foundations of Newtonian mechanics. This great book, which showed for the first time how earthly and heavenly movements obey the same laws, is cast in the form of a set of propositions all deriving from three axioms, or *laws of motion*. It is these, here translated into modern English from the original Latin, that still provide the basis for Newtonian mechanics.

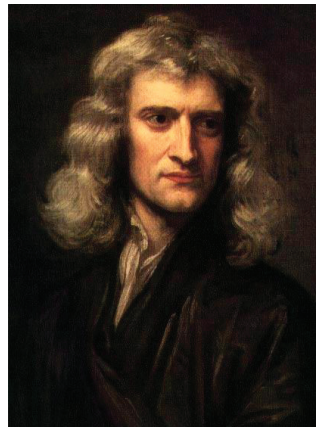


Figure 1 Sir Isaac Newton (1642–1727)

Law I Every body continues in a state of rest, or moves with constant velocity in a straight line, unless a force is applied to it.

Law II The rate of change of velocity of a body is proportional to the resultant force applied to the body, and happens in the direction of the resultant force.

Law III To every action (i.e. force) by one body on another there is always opposed an equal reaction (i.e. force) – that is, the actions of two bodies on each other are always equal in magnitude and opposite in direction.

One of the central concepts in Newtonian mechanics is that of a *force*. The word 'force' is used in everyday conversation in a variety of ways: he *forced* his way in; the *force* of destiny; to put into *force*; the labour *force*. In mathematics and science, the word 'force' has a precise definition. However, this definition relies on the movement of objects thus is deferred until the next unit. Essentially, though, this definition says that a **force** either changes the shape of the object on which it acts, or causes movement of the object. When we experience a force, in the mathematical sense of the word, we feel it through contact: pulling on a rope, lifting a shopping bag, pushing against a car, holding a child aloft. In each case, the force that we experience has a magnitude and a direction, so we model a force as a *vector quantity*.

We often need to represent physical quantities – such as mass, force, velocity, acceleration, time – mathematically. Most of the physical quantities that we need can be classified into two types: *scalars* and *vectors*. *Scalar quantities* are quantities, like mass, temperature, energy, volume and time, that can be represented by a single real number. Other quantities, like force, velocity and acceleration, possess *magnitude* and *direction* in space, and cannot be represented by a single real number; they are called *vector quantities*.

The definitive vector quantity is *displacement*. The displacement of a point specifies the position of the point in space relative to some reference point. We use the concept of displacement whenever we want to describe spatial relationships.

Section 1 defines a vector and discusses ways of representing vectors in two dimensions. Section 2 discusses another way of representing vectors, one that easily generalises from two to three (or more) dimensions, and considers ways of operating on and combining vectors – that is, the fundamentals of *vector algebra*.

Mathematical representation of a force

A force is represented mathematically by a vector. The magnitude of the vector represents the magnitude of the force, and the direction of the vector specifies the direction in which the force is applied.

Sometimes we can see the effect of a force, such as when a mattress depresses under the weight of someone sitting on it, a washing line sags under the weight of the washing, a rubber band is stretched over some packages, a door is pulled open, or a bag of shopping is lifted. In each case there is an obvious deformation or movement that indicates that a force is present. Sometimes, however, the presence of a force is not so obvious, for example in situations such as a ladder leaning against a wall (though you could appreciate the presence of a force in this situation were *you* to replace the wall and hold the ladder steady yourself), a box resting on a shelf, or a cable holding up a ceiling lamp (consider holding up the lamp by the cable yourself).

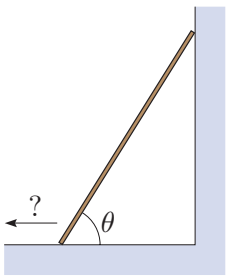


Figure 2 A ladder leaning against a wall

The second half of this unit considers the conditions under which objects remain stationary when subjected to forces, which is a topic known as **statics**. For example, what is the minimum angle θ that a ladder leaning against a wall can make with the ground before the ladder slides to the ground (see Figure 2)? Cases where forces cause motion are discussed in Unit 3 and elsewhere in the module. The study of motion is called **dynamics**.

Before forces and their effects can be analysed mathematically, they and the objects on which they act must be modelled mathematically. In Section 3, objects are modelled as particles, and we show how various forces such as the forces of gravity, tension and friction can be modelled. We also show how to analyse one-particle systems in equilibrium, and extend the ideas to systems involving two or more particles. Section 4 considers situations where an object needs to be modelled as a solid body rather than as a particle. It also discusses the turning effect of a force (known as a *torque*), which happens only if a force acts on a solid body rather than a particle. Section 5 describes the application to statics problems of the concepts and principles described in the earlier sections.

1 Describing, representing and combining vectors

Subsection 1.1 explains what scalars and vectors are. Subsection 1.2 explains how to denote vectors symbolically (i.e. algebraically) and how to show them in diagrams. It then explains what is meant by saying that two vectors are equal to one another, which is a necessary first step in the development of an algebra for vectors. Subsection 1.3 describes how a vector can be scaled by multiplying it by a scalar, and how vectors can be added together.

1.1 Scalars and vectors

A **scalar** is any quantity, such as mass, time, volume or temperature, that can be represented mathematically by a single real number (and often a unit of measurement). Real numbers themselves are examples of scalars, and you can regard the terms *scalar* and *real number* as synonymous. Examples of scalar quantities, quoted to some convenient degree of accuracy, are:

- the mass of the Earth, 5.975×10^{24} kilograms
- the temperature of melting ice, 0 degrees Celsius
- my current bank account balance, -153.12 pounds sterling
- pi (π), 3.141 592 653 589 79....

A real number x is defined by two properties: its *modulus* $|x|$ and its *sign*. Thus the **magnitude** of a scalar x is $|x|$. For example, the magnitude of my current account balance is $|-153.12|$ pounds $= 153.12$ pounds, which sounds a lot better since it doesn't remind me that I'm in debt. Note that magnitudes are always non-negative (i.e. positive or zero).

The modulus of a real number is also called its *magnitude*.

A **vector** is any quantity, such as force, velocity or displacement, that has a magnitude *and* a direction in space (or, in two dimensions, a direction in a plane). An example is the velocity of a motor car travelling on the M4 motorway from London to Bristol with a speed of 95 km per hour in a westerly direction. The magnitude of the velocity vector is 95 (dropping units for convenience), and the direction of the velocity vector is due west. Thus the specification of a vector consists of:

The familiar term *speed* is used to mean the magnitude of velocity. Speed is a non-negative scalar.

- a non-negative real number, called its modulus or **magnitude**
- a **direction** in space.

Vectors are denoted in printed text by bold letters, e.g. $\mathbf{v}$, $\mathbf{F}$. In handwritten work, you should denote vectors by drawing either a straight line or a wavy line under the letter, e.g. $\underline{\mathrm{v}}$, $\underline{\mathrm{F}}$ or $\underset{\sim}{\mathrm{v}}$, $\underset{\sim}{\mathrm{F}}$. Thus if a symbol is used to represent the velocity of an object, then it must be handwritten by you as either $\underline{\mathrm{v}}$ or $\underset{\sim}{\mathrm{v}}$ (but will be printed in the text as $\mathbf{v}$). An exception to this rule is that a vector representing the displacement of the point Q from the point P is often written as $\overrightarrow{PQ}$, and if $\mathbf{x}$ represents this displacement vector, then we can write $\mathbf{x} = \overrightarrow{PQ}$.

It is *important* that you learn to write vectors using underlining: if you do not do so, then someone reading your work may not be able to tell that you are referring to a vector. In particular, you may lose marks in assessed work!

We read $|\mathbf{v}|$ as 'the modulus of v' or 'the magnitude of v', or simply 'mod v'.

The modulus or magnitude of the vector $\mathbf{v}$ is denoted by $|\mathbf{v}|$, or sometimes, where there is no possibility of ambiguity, by v; $|\mathbf{v}|$ is a non-negative scalar.

1.2 Using arrows to represent vectors

A vector can be conveniently represented in a diagram by an arrow, that is, a straight line with an arrowhead on it. The tail of the arrow may be placed at some fixed origin, its direction is chosen to represent that of the vector, and its length is chosen to be proportional to the magnitude of the vector. In Figure 3, which uses the origin of the Cartesian coordinate system as the fixed origin, the shorter arrow represents a vector of magnitude 1 in the positive x-direction, and the longer arrow represents a vector of magnitude $2\sqrt{2}$ in a direction at $\pi/4$ radians (45°) to the positive x-direction. (Note that we use the convention that positive angles are measured anticlockwise.) If we decide to denote these vectors by letters $\mathbf{a}$ and $\mathbf{b}$, then we can also put this information on the diagram, by writing $\mathbf{a}$ and $\mathbf{b}$ near the arrowheads.

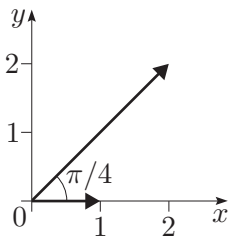


Figure 3 Examples of vectors

Exercise 1

Represent the following two vectors on a diagram by arrows:

- vector $\mathbf{a}$ has magnitude 3 units and points in the positive y-direction
- vector $\mathbf{b}$ has magnitude 4 units and points in the direction at $\pi/3$ radians (60°) to the positive x-direction.

What about a vector whose magnitude is zero? Clearly its length is zero, but what is its direction? The answer is that it does not have one! We make the following definition.

Zero vector

The **zero vector** is the unique vector with magnitude zero and no direction. It is denoted by $\mathbf{0}$.

Be particularly careful to underline the zero vector ($\underline{0}$ or $\underset{\sim}{0}$) in your written work, and be aware that the vector $\mathbf{0}$ is different from the scalar 0!

A constant *velocity* is defined by a magnitude and a direction. For instance, in a weather forecast, a typical wind velocity might be 35 mph from the north-west. It is not sufficient to say that 'the wind velocity is 35 mph'; the obvious question about such a statement would be 'from which direction?'. The vector $\mathbf{v}$ representing this velocity has magnitude 35 and direction from the north-west and towards the south-east (since the air is travelling in the south-easterly direction).

It can be represented on a diagram like Figure 4. The length of the arrow represents a wind speed of 35 mph.

Note that the *direction* of a vector consists of two attributes:

- an *orientation*, represented by the *slope* of the arrow in diagrams like Figures 3 and 4
- a *sense*, represented by the *arrowhead*.

For instance, the arrow representing the velocity 35 mph from the north-west in Figure 4 is a line making an angle of 45° anticlockwise from the south direction (the orientation) and an arrowhead pointing towards south-east as opposed to north-west (the sense).

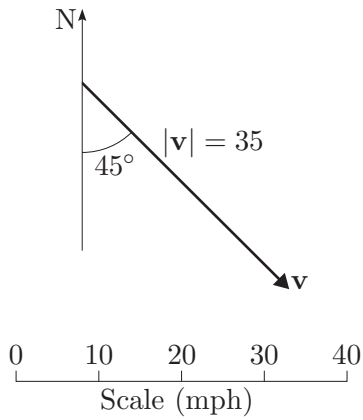


Figure 4 A vector representing wind speed and direction

Exercise 2

A car travelling from London along the M1 with speed 70 mph heads in the direction N 60° W near Junction 14.

Represent the velocity of the car by an arrow, drawn to a suitable scale.

You have seen how to represent a vector by an arrow. Here is a definition of equality of vectors.

> Two vectors are said to be **equal** if they have the same magnitude *and* the same direction.

This definition tells us that the two features needed to define a vector uniquely are its magnitude and direction. This means that any two arrows that are drawn at different places on the page but are equal in length, are parallel and have the same sense, can be used to represent the same vector. For instance, the two arrows in Figure 5 are each of length 2 units and point in the positive x-direction. They represent two equal vectors, and we write $\mathbf{b} = \mathbf{d}$. In other words, the arrow representing a vector does not have to be drawn so that its tail is at any particular point.

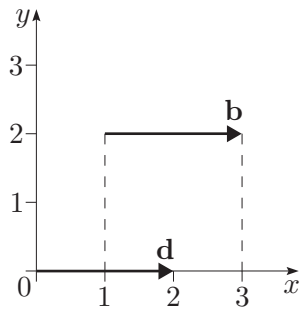


Figure 5 Two equal vectors

1.3 Vector algebra

If $\mathbf{v}$ is a vector and m is a positive number, then the product $m\mathbf{v}$ is a vector in the same direction as $\mathbf{v}$ but with magnitude $m|\mathbf{v}|$, that is, m times the magnitude of $\mathbf{v}$. This multiplication of a vector by a scalar is called *scaling* or *scalar multiplication*, and $m\mathbf{v}$ is called a *scalar multiple* of $\mathbf{v}$. For example, if $\mathbf{v}$ has magnitude 4 and points in the positive x-direction, then $3\mathbf{v}$ has magnitude 12 and points in the positive x-direction also. This is illustrated in Figure 6.

Note that there is no multiplication sign between the m and the $\mathbf{v}$. In vector algebra, the dot and cross symbols are reserved for other products, to be discussed in Section 2.

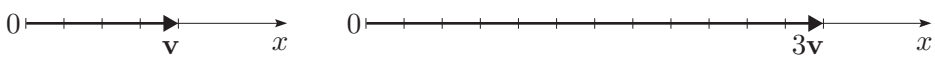


Figure 6 Scaling a vector

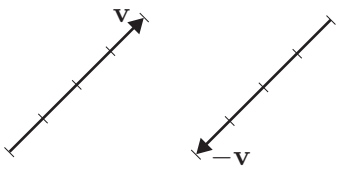


Figure 7 Reversing the sign of a vector

We can also scale a vector $\mathbf{v}$ by a negative number. When m is negative, the vector $m\mathbf{v}$ has magnitude $|m|\,|\mathbf{v}|$ and points in the *opposite* direction to $\mathbf{v}$. A special case is when $m = -1$. Then the vector $(-1)\mathbf{v}$ has the same magnitude as $\mathbf{v}$ but points in the opposite direction; see Figure 7. We normally write $(-1)\mathbf{v}$ simply as $-\mathbf{v}$, that is, $(-1)\mathbf{v} = -\mathbf{v}$.

For any vector $\mathbf{v}$ and any real number m, the **scalar multiple** $m\mathbf{v}$ is the vector with magnitude $|m|\,|\mathbf{v}|$ that is:

- in the same direction as $\mathbf{v}$ if $m > 0$
- in the opposite direction to $\mathbf{v}$ if $m < 0$
- the zero vector (i.e. with unspecified direction) if $m = 0$.

The multiplication of $\mathbf{v}$ by m is called **scaling** or **scalar multiplication**.

Exercise 3

(a) If $\mathbf{v}$ represents the velocity of a wind of 35 mph from the north-east, what vector represents a wind of 35 mph from the south-west?

(b) Relate the direction and magnitude of $-1.5\mathbf{v}$ to those of $\mathbf{v}$, where $\mathbf{v}$ is any given non-zero vector. Do the same for $-k\mathbf{v}$, where k is an arbitrary positive number.

(c) If $\mathbf{v}$ is any non-zero vector, what are the magnitude and direction of the vector $\dfrac{1}{|\mathbf{v}|}\mathbf{v}$?

The vector $\dfrac{1}{|\mathbf{v}|}\mathbf{v}$ in Exercise 3(c) is a vector that has magnitude 1 and points in the direction of $\mathbf{v}$. It is called the *unit vector* in the direction of $\mathbf{v}$ and is often denoted by the symbol $\widehat{\mathbf{v}}$ (read as 'v hat').

For any non-zero vector $\mathbf{v}$, the **unit vector** in the direction of $\mathbf{v}$ is the vector

$$\widehat{\mathbf{v}} = \frac{1}{|\mathbf{v}|}\mathbf{v}. \qquad (1)$$

Unit vectors are often used to denote directions in the plane or in space.

A particular example is provided by the unit vectors in the positive directions of the x- and y-axes in the plane Cartesian coordinate system. (We will develop the Cartesian representation of vectors in Section 2.) These unit vectors are denoted by $\mathbf{i}$ and $\mathbf{j}$, respectively, and are called **Cartesian unit vectors**. Note that we do not denote these unit vectors as $\widehat{\mathbf{i}}$ and $\widehat{\mathbf{j}}$ as they are widely used without the 'hat' in mathematics.

Figure 8 shows these Cartesian unit vectors and two other vectors, **a** and **b**. The vector **a** has magnitude 2 and points in the positive x-direction; **b** has magnitude 3.5 and points in the positive y-direction. The unit vector **i** has magnitude 1 and points in the same direction as **a**. Thus we can write **a** in terms of **i** by a scaling:

$$\mathbf{a} = 2\,\mathbf{i}.$$

Similarly, we can write **b** in terms of **j**:

$$\mathbf{b} = 3.5\,\mathbf{j}.$$

Any vector parallel to the x- or y-axis can be written as a scaling of **i** or **j**.

Note that although **i** and **j** are shown in Figure 8 with their tails at the origin, this is not necessary. They can be drawn at any convenient position, provided only that they are of unit magnitude and point in the positive x- and y-directions, respectively.

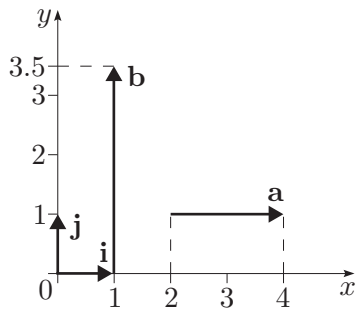


Figure 8 Vectors **a** and **b**, and two unit vectors, **i** and **j**

Exercise 4

Let the unit vectors **i** and **j** denote the directions of east and north, respectively. Specify the following vectors as scalings of **i** and **j**.

(a) A wind velocity of 35 km per hour due south

(b) The displacement of Bristol from London (112 miles due west)

Let us consider what is meant by the addition of vectors. Suppose that we make a journey from Bristol to Leeds, and then another journey from Leeds to Norwich. The first journey produces a displacement of $\mathbf{d}_1$ and the second a displacement of $\mathbf{d}_2$. The net result of the two journeys is a displacement of $\mathbf{d}_3$ from Bristol to Norwich. This is illustrated by the triangle of displacements shown in Figure 9. Displacements are said to add by the *triangle rule*, and we write $\mathbf{d}_3 = \mathbf{d}_1 + \mathbf{d}_2$. The vector $\mathbf{d}_3$ is called the *resultant* of $\mathbf{d}_1$ and $\mathbf{d}_2$.

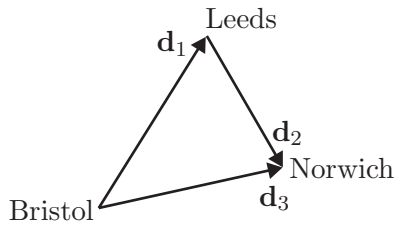


Figure 9 A triangle of displacements

Velocities also add by the triangle rule, and so do forces, accelerations and all other vector quantities. Thus the triangle rule is also called the *vector addition rule*.

Triangle rule or vector addition rule

To add any two vectors **a** and **b**: choose an origin O; draw the line OP in the direction of **a** and with length equal to the magnitude of **a**; and draw the line PQ in the direction of **b** and with length equal to the magnitude of **b** (as in Figure 10). Then $\mathbf{a} + \mathbf{b}$ is the vector with magnitude equal to the length of OQ and with direction from O to Q. The vector $\mathbf{a} + \mathbf{b}$ is called the **sum** or **resultant** of **a** and **b**.

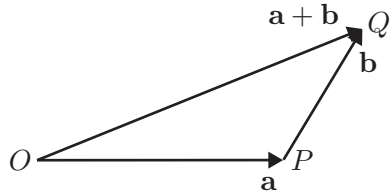


Figure 10 The triangle rule or vector addition rule

Now recall that when discussing displacements, we mentioned the zero vector **0** (representing no displacement). Once addition of vectors is introduced, we *need* the zero vector in order to answer questions such as 'What is $\mathbf{i} + (-1)\mathbf{i}$?'. Geometrically, no construction is needed when adding the zero vector, which obeys the rather obvious rule

$$\mathbf{a} + \mathbf{0} = \mathbf{a}.$$

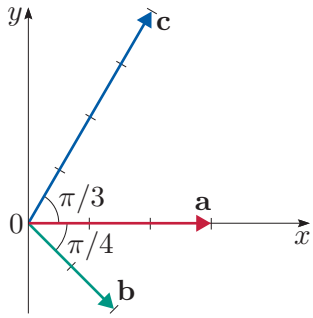


Figure 11

Exercise 5

Three vectors **a**, **b** and **c** of magnitudes 3, 2 and 4 are shown in Figure 11.

(a) Draw a rough sketch to show the vectors $\mathbf{a} + \mathbf{b}$ and $\mathbf{a} + \mathbf{c}$.

(b) Sketch the vector $-\mathbf{b}$, and draw a rough sketch to show the sum of **a** and $-\mathbf{b}$.

Exercise 5(b) suggests a definition of **vector subtraction**. To subtract the vector **b** from the vector **a**, we *add* the vectors **a** and $-\mathbf{b}$ by the triangle rule of vector addition; that is, in symbols,

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b}).$$

Exercise 6

A vector **a** has magnitude 3 units and points in the positive x-direction. A vector **b** has magnitude 4 units and points in the positive y-direction. Draw a diagram showing the vectors $\mathbf{a} + \mathbf{b}$ and $\mathbf{a} - \mathbf{b}$.

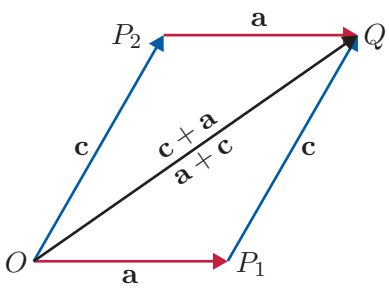


Figure 12 The parallelogram rule

Vector addition is *commutative*, that is, the order in which we add two vectors does not matter. This can be illustrated by reference to vectors **a** and **c** of Exercise 5 (see Figure 12). The triangle OP_1Q illustrates the addition $\mathbf{a} + \mathbf{c}$, while triangle OP_2Q illustrates $\mathbf{c} + \mathbf{a}$. The same resultant $\overrightarrow{OQ}$ is obtained in both cases. Thus

$$\mathbf{a} + \mathbf{c} = \mathbf{c} + \mathbf{a}.$$

An alternative geometric construction for adding two vectors can be seen from Figure 12. It is called the **parallelogram rule**. Draw the two vectors **a** and **b** with the same beginning point O. Complete the parallelogram OP_1QP_2. Then the resultant vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$ is the vector on the diagonal of the parallelogram. The parallelogram rule gives the same resultant as the triangle rule.

The algebraic rules are summarised below.

Algebraic rules for scaling and adding vectors

- A vector has magnitude and direction.
- Two vectors can be added by the triangle rule.
- A vector can be scaled by a real number. For example, any non-zero vector $\mathbf{a}$ can be written as $\mathbf{a} = |\mathbf{a}|\,\widehat{\mathbf{a}}$, where $|\mathbf{a}|$ is the magnitude of $\mathbf{a}$, and $\widehat{\mathbf{a}}$ is a unit vector in the direction of $\mathbf{a}$.

Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be vectors, and let m, m_1 and m_2 be scalars.

- Addition is commutative: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$.
- Addition is associative: $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$.
- $m\mathbf{a}$ is a vector with magnitude $|m|\,|\mathbf{a}|$, in the same direction as $\mathbf{a}$ when $m > 0$, and in the opposite direction when $m < 0$.
- Scaling is associative: $m_1(m_2\mathbf{a}) = (m_1 m_2)\mathbf{a}$.
- Scaling is distributive: $(m_1 + m_2)\mathbf{a} = m_1\mathbf{a} + m_2\mathbf{a}$.
- Scaling is distributive over vector addition: $m(\mathbf{a} + \mathbf{b}) = m\mathbf{a} + m\mathbf{b}$.
- Addition and scaling involving the zero vector are as expected: $\mathbf{0} + \mathbf{a} = \mathbf{a}$ and $0\mathbf{a} = \mathbf{0}$.
- Subtraction is defined by $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-1)\mathbf{b}$.

These rules allow us to manipulate algebraic expressions involving scalings and vector addition in a familiar way.

Example 1

Simplify the expression $2(\mathbf{a} + \mathbf{b}) + 3(\mathbf{b} + \mathbf{c}) - 5(\mathbf{a} + \mathbf{b} - \mathbf{c})$.

Solution

We have

$$\begin{aligned} &2(\mathbf{a} + \mathbf{b}) + 3(\mathbf{b} + \mathbf{c}) - 5(\mathbf{a} + \mathbf{b} - \mathbf{c}) \\ &\quad = 2\mathbf{a} + 2\mathbf{b} + 3\mathbf{b} + 3\mathbf{c} - 5\mathbf{a} - 5\mathbf{b} + 5\mathbf{c} \\ &\quad = 8\mathbf{c} - 3\mathbf{a}. \end{aligned}$$

Exercise 7

Simplify the expression $4(\mathbf{a} - \mathbf{c}) + 3(\mathbf{c} - \mathbf{b}) + 2(2\mathbf{a} - \mathbf{b} - 3\mathbf{c})$.

2 Cartesian components and products of vectors

So far we have approached vectors, and the rules of vector addition and scaling, geometrically. To add vectors geometrically requires drawing diagrams representing the vectors by arrows. An alternative, and sometimes more convenient, *algebraic* approach to representing three-dimensional vectors is developed in this section.

So far in this unit we have defined two algebraic operations: vector addition (by the triangle rule) and scaling a vector. The addition of vectors can be usefully applied only to two vectors representing the same type of physical quantity. For example, the addition of a displacement and a velocity has no physical meaning. However, vectors representing the same or different types of physical quantity can be combined in operations that are called the *dot product* and the *cross product*. They are called products because in some respects they behave like 'multiplications' in the algebra of real numbers. Dot products and cross products of vectors have numerous applications in geometry, mechanics and electromagnetism.

In Subsections 2.2 and 2.3 the dot product and cross product are defined geometrically and also in terms of components of vectors. The dot product of two vectors is interpreted in terms of projecting a shadow of one vector onto another, and is applied to the problem of finding the angle between two vectors or lines. The cross product of two vectors is interpreted as a vector whose magnitude is an area. Both dot and cross products can be used in problems involving finding the areas of plane figures and the volumes of solid objects.

2.1 Vectors in three dimensions

In previous modules you may have discussed vectors in the plane. However, the world is three-dimensional, and few real problems are restricted to a plane surface. For example, starting at point A at one corner of the cube shown in Figure 13, you can reach the opposite corner S by three successive displacements: $\overrightarrow{AQ} + \overrightarrow{QB} + \overrightarrow{BS}$. In order to work with such addition of displacements in three dimensions, it is necessary to introduce a three-dimensional coordinate system.

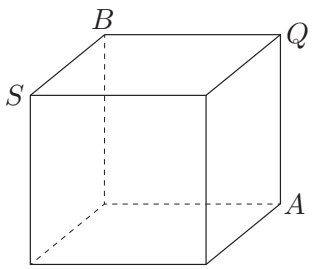


Figure 13 A cube

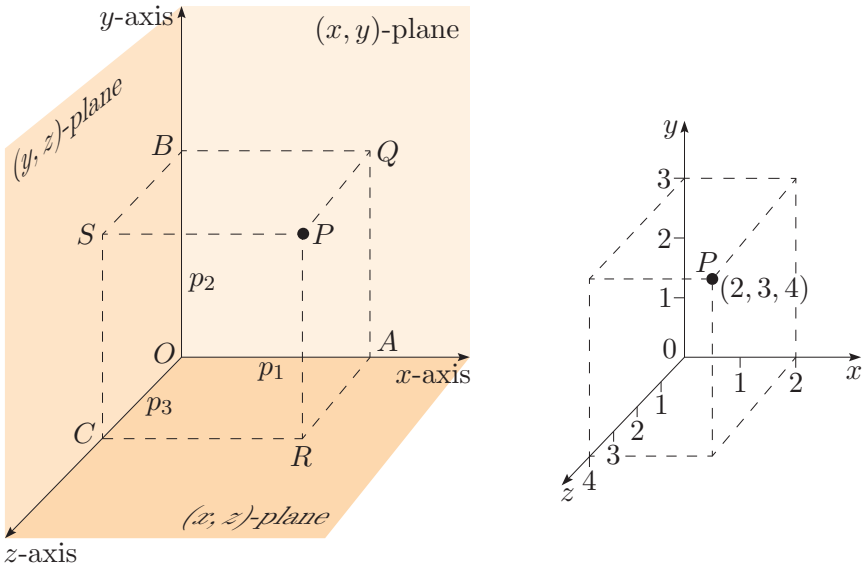
A three-dimensional Cartesian coordinate system

Consider a two-dimensional Cartesian coordinate system Oxy. Draw a third axis, the z-axis, through the origin O, perpendicular to both the x- and y-axes of the two-dimensional system. This produces a coordinate system with three mutually perpendicular axes, the x-, y- and z-axes (see Figure 14), intersecting at O.

The z-axes shown in Figure 14 are meant to point *out* of the plane of the page.

Alternatively, the coordinate system can be characterised by three planes:

- the **(x, y)-plane**, which contains the x- and y-axes and is perpendicular to the z-axis
- the **(x, z)-plane**, analogously defined
- the **(y, z)-plane**, again analogously defined.



(a) (b)

Figure 14 The Cartesian coordinates of a point in space

Any point P can be represented uniquely by its perpendicular distances from the (x, y)-, (x, z)- and (y, z)-planes. These distances, called the (**Cartesian**) **coordinates** of P, are shown in Figure 14(a). QP, RP and SP are perpendicular to the (x, y)-plane, (x, z)-plane and (y, z)-plane, respectively.

We denote the point P by the ordered triple of coordinates (p_1, p_2, p_3), where

$$p_1 = SP = OA,$$
$$p_2 = RP = OB,$$
$$p_3 = QP = OC.$$

For example, the point $(2, 3, 4)$ is shown in Figure 14(b).

When drawing Figure 14 it was necessary to choose one of two possible ways for the positive z-direction to be defined; these are shown in Figure 15, where the z-axis is meant to point out of the plane of the page in the top diagram, and into the plane of the page in the bottom diagram.

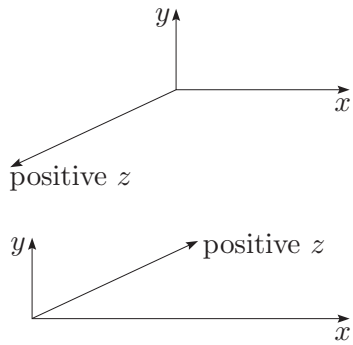


Figure 15 Choosing the direction of positive z

The usual convention for relating the positive directions of x, y and z is given by the following rule, called the **right-hand rule**. The right hand is held with the middle finger, first finger and thumb placed (roughly) perpendicular to each other, and the other two fingers closed (see Figure 16). If the thumb and first finger are pointing in the directions of the positive x- and y-axes, respectively, then the middle finger is pointing in the direction of the positive z-axis.

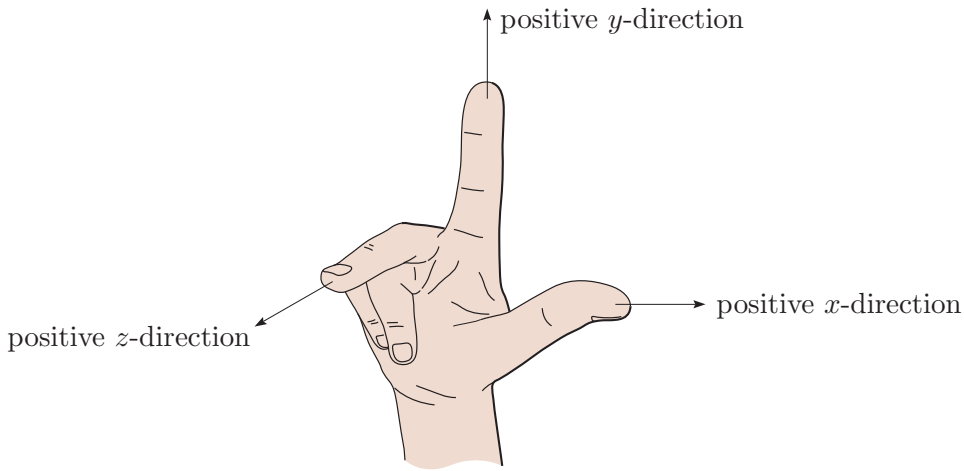


Figure 16 The right-hand rule

Alternatively, you can think of Figure 14(a) as showing a corner of a room (with the y-axis pointing upwards). If you are standing in the corner facing outwards, then the left-hand edge of the floor is the x-axis, and the right-hand edge is the z-axis. A coordinate system defined in this way is called a **right-handed system**. Only right-handed systems will be used in this module. The systems drawn in Figure 14 and the top of Figure 15 are right-handed systems.

An alternative definition of the same positive z-direction is given by the **right-hand grip rule**, stated as follows (see Figure 17). If you hold your right hand in a fist, so that your fingers rotate from the positive x-axis towards the positive y-axis, then your thumb points in the positive z-direction of a right-handed coordinate system.

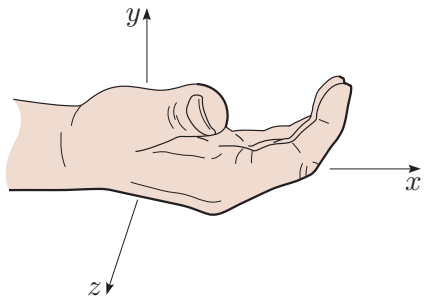


Figure 17 The right-hand grip rule

You should use whichever rule you find easier to apply.

Exercise 8

Decide which of the sets of perpendicular axes in Figure 18 define right-handed coordinate systems.

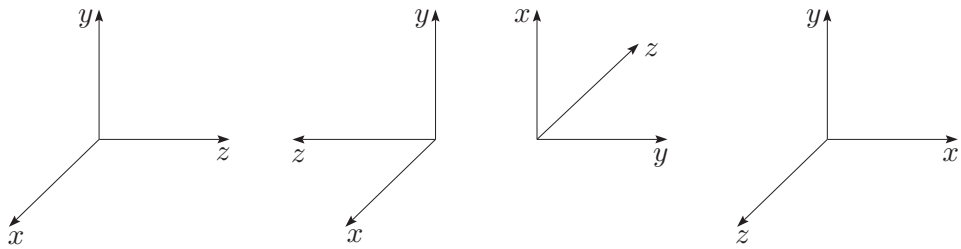


Figure 18

(The x-axis points out of the plane of the page in (a) and (b). The z-axis points into and out of the plane of the page in (c) and (d), respectively.)

Component form of three-dimensional vectors

The algebraic representation of vectors can be extended to vectors in three dimensions, such as in Figure 19. The vector $\mathbf{a}$, drawn from the origin O, is the *position vector* of point A with three-dimensional Cartesian coordinates (a_1, a_2, a_3). A third Cartesian unit vector $\mathbf{k}$ is introduced to represent the positive z-direction. We now have three Cartesian unit vectors, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, which are perpendicular to each other. The vector $\mathbf{a}$ may thus be written in *component form* as

$$\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k} \quad \text{or} \quad \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{or} \quad \mathbf{a} = (a_1 \quad a_2 \quad a_3)^T. \qquad (2)$$

The third form above is often used in the text, to save space. The *transpose symbol* T indicates that a column vector has been written 'on its side' as a row vector.

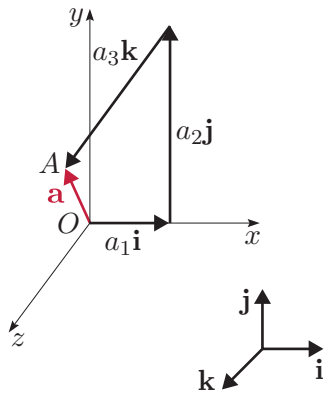


Figure 19 Expressing a vector in component form

The **position vector** of a point A relative to the origin O of three-dimensional space is the displacement of A from O, that is, the vector

$$\mathbf{a} = \overrightarrow{OA}.$$

The **i**-, **j**- and **k**-components of the position vector $\mathbf{a}$ are the coordinates a_1, a_2 and a_3 of the point A, respectively.

These may sometimes be referred to as x-, y- and z-components.

The components of vectors not based at the origin are defined similarly, as follows.

A vector $\mathbf{a} = \overrightarrow{PQ}$ in three-dimensional space, where P is the point (p_1, p_2, p_3) and Q is the point (q_1, q_2, q_3), has **component form**

$$\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}, \tag{3}$$

where $a_1 = q_1 - p_1$, $a_2 = q_2 - p_2$, $a_3 = q_3 - p_3$, and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are the Cartesian unit vectors. The numbers a_1, a_2, a_3 are the (**Cartesian**) **components** of $\mathbf{a}$.

Note that this component form may also be written as in (2).

For example, in Figure 14(b), we have $\overrightarrow{OP} = 2\,\mathbf{i} + 3\,\mathbf{j} + 4\,\mathbf{k}$.

As in two dimensions, the operations of vector algebra can be expressed in terms of components.

Adding and scaling three-dimensional vectors in component form

If $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$, $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$ and m is a scalar, then

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k} \tag{4}$$

and

$$m\mathbf{a} = (ma_1)\mathbf{i} + (ma_2)\mathbf{j} + (ma_3)\mathbf{k}. \tag{5}$$

The magnitude of a vector in terms of its components a_1, a_2, a_3 can be found using Pythagoras' theorem (see Figure 20).

The length ON is $\sqrt{a_1^2 + a_2^2}$, and $OA^2 = ON^2 + NA^2$. But $OA = |\mathbf{a}|$, thus

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

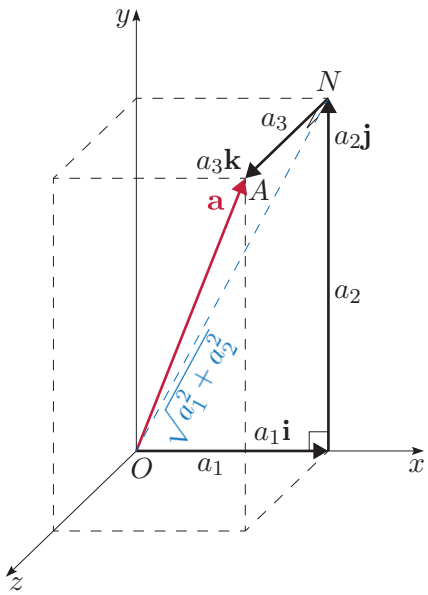


Figure 20 Finding the magnitude of a vector

This can be summarised as follows.

Magnitude of a three-dimensional vector in component form

If $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$, then its magnitude is

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{6}$$

A unit vector in the direction of $\mathbf{a}$ is

$$\widehat{\mathbf{a}} = \frac{a_1}{\sqrt{a_1^2 + a_2^2 + a_3^2}}\,\mathbf{i} + \frac{a_2}{\sqrt{a_1^2 + a_2^2 + a_3^2}}\,\mathbf{j} + \frac{a_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}}\,\mathbf{k}. \tag{7}$$

Exercise 9

Consider the vectors $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - 3\mathbf{j} - \mathbf{k}$ and $\mathbf{c} = 3\mathbf{i} + \mathbf{k}$.

(a) Express $\mathbf{d} = 2\mathbf{a} - 3\mathbf{b}$ and $\mathbf{e} = \mathbf{a} - 2\mathbf{b} + 4\mathbf{c}$ in component form.

(b) Find the magnitudes of the vectors $\mathbf{d}$ and $\mathbf{e}$.

(c) Evaluate $|\mathbf{a}|$, and write down a unit vector in the direction of $\mathbf{a}$.

(d) Find the components of a vector $\mathbf{x}$ such that $\mathbf{a} + \mathbf{x} = \mathbf{b}$.

Exercise 10

Write the vectors $\mathbf{0}$, $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as column vectors in three dimensions.

Vector equation of a straight line

One useful application of position vectors (in two or three dimensions) is in obtaining a vector equation of a straight line.

Example 2

Find the position vector of a point P lying on a straight-line segment AB in terms of the position vectors of A and B.

Solution

Let P be any point on AB (see Figure 21). The position vector $\overrightarrow{OP}$ of P relative to the origin can also be written, using the triangle rule, as

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AP}.$$

Now $\overrightarrow{AP} = s\,\overrightarrow{AB}$, for some number s, and the point P traces out the line segment AB as s varies from 0 to 1. Thus the straight-line segment AB is described by the vector equation

$$\overrightarrow{OP} = \overrightarrow{OA} + s\,\overrightarrow{AB} \quad (0 \leq s \leq 1).$$

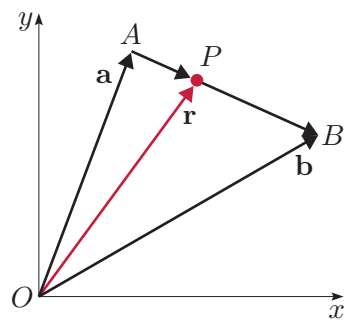


Figure 21 Finding a general point on a straight line

Writing $\mathbf{a} = \overrightarrow{OA}$, $\mathbf{b} = \overrightarrow{OB}$, $\mathbf{r} = \overrightarrow{OP}$, and noting (using the triangle rule) that $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \mathbf{b} - \mathbf{a}$, this equation can also be written as

$$\mathbf{r} = \mathbf{a} + s(\mathbf{b} - \mathbf{a}) = (1 - s)\mathbf{a} + s\mathbf{b} \quad (0 \leq s \leq 1).$$

Note that if the parameter s in Example 2 is allowed to range over all the real numbers ($-\infty < s < \infty$), then the point P traces out the entire straight line of which AB is a segment. Also note that the ideas in Example 2 are easily extended to three dimensions.

Vector equation of a straight line

If A and B are any two distinct points on a straight line in space, with position vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, with respect to some given origin, then the **vector equation of the straight line** is

$$\mathbf{r} = (1 - s)\mathbf{a} + s\mathbf{b} \quad (-\infty < s < \infty), \tag{8}$$

where $\mathbf{r}$ represents the position vector of any point on the line.

If $0 \leq s \leq 1$, then the equation represents only the line segment AB.

In Cartesian coordinate form, with $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$, $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$ and $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, the equation of the line is

$$x = a_1 + s(b_1 - a_1), \quad y = a_2 + s(b_2 - a_2), \quad z = a_3 + s(b_3 - a_3).$$

This is called the **parametric form of the straight line**, where the coordinates are expressed in terms of the parameter s. We can recover a more familiar form of the straight line by writing everything in terms of s as

$$\frac{x - a_1}{b_1 - a_1} = \frac{y - a_2}{b_2 - a_2} = \frac{z - a_3}{b_3 - a_3} (= s).$$

Exercise 11

(a) Write down, in component form, the vector equation of the straight line on which lie the points with Cartesian coordinates $(1, 1, 2)$ and $(2, 3, 1)$.

(b) Find the coordinates of the point where the line cuts the (x, z)-plane.

Exercise 12

Let $\mathbf{a} = 2\mathbf{i} - \mathbf{j}$, $\mathbf{b} = \mathbf{i} + 3\mathbf{j} + 5\mathbf{k}$ and $\mathbf{c} = \mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

(a) Find the magnitudes of $\mathbf{a}$ and $\mathbf{b}$, and describe the direction of $\mathbf{a}$.

(b) Find the vectors $\mathbf{a} + \mathbf{b}$, $2\mathbf{a} - \mathbf{b}$ and $\mathbf{c} + 2\mathbf{b} - 3\mathbf{a}$ in component form.

(c) What is the endpoint Q of the displacement represented by the vector $2\mathbf{a} - \mathbf{b}$ if $(0, 2, 3)$ is its beginning point P?

Exercise 13

(a) A straight line in Cartesian form is expressed as

$$\frac{x+1}{3} = \frac{y-8}{-2} = 1 - z.$$

Express this equation in parametric and vector form. If A is the point corresponding to $s = 0$, and B is the point corresponding to $s = 1$, determine the component form of $\overrightarrow{AB}$ and the distance between A and B.

(b) (i) Determine the vector position of M, the point midway between A and B.

(ii) Determine the vector position of N, the point that divides the line AB in the ratio $3 : 2$.

(c) Find the vector equation of the straight line through the points $(-1, -2, 0)$ and $(2, 1, -1)$.

(d) Determine whether the lines in parts (a) and (c) intersect. If so, determine the coordinates of the point of intersection.

2.2 The dot product

We make the following definition.

The **dot product** of two vectors $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\theta, \tag{9}$$

where θ $(0 \leq \theta \leq \pi)$ is the angle between the directions of $\mathbf{a}$ and $\mathbf{b}$ (see Figure 22).

The product $\mathbf{a} \cdot \mathbf{b}$ is read as 'a dot b'.

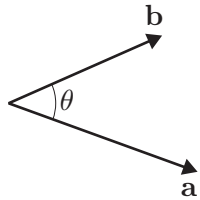


Figure 22 The angle between two vectors

The dot product of two vectors is a *scalar quantity*, that is, it is a real number: $\mathbf{a} \cdot \mathbf{b}$ is the product of the three scalars $|\mathbf{a}|$, $|\mathbf{b}|$ and $\cos\theta$. So the operation of the dot product combines two vectors to define a scalar, and for this reason the dot product is also called the **scalar product**. The angle θ lies in the range $0 \leq \theta \leq \pi$, and the value of $\mathbf{a} \cdot \mathbf{b}$ is:

- positive for $0 \leq \theta < \frac{\pi}{2}$, i.e. when θ is an acute angle
- negative for $\frac{\pi}{2} < \theta \leq \pi$, i.e. when θ is obtuse
- zero for $\theta = \frac{\pi}{2}$, i.e. when θ is a right angle.

It is important, when writing a dot product, to make sure that the dot between the vectors is clear.

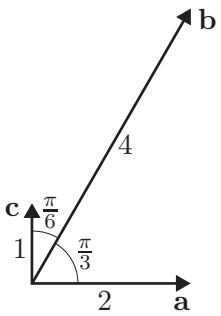


Figure 23

Exercise 14

Three vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ of magnitudes 2, 4 and 1 units, respectively, lying in the same plane, are represented by arrows as shown in Figure 23. The angle between the vectors $\mathbf{a}$ and $\mathbf{b}$ is $\frac{\pi}{3}$ radians, and that between the vectors $\mathbf{b}$ and $\mathbf{c}$ is $\frac{\pi}{6}$ radians. Use the definition of the dot product to find the values of $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{b} \cdot \mathbf{c}$, $\mathbf{a} \cdot \mathbf{c}$ and $\mathbf{b} \cdot \mathbf{b}$.

This exercise demonstrates two important properties of the dot product.

- If two vectors $\mathbf{a}$ and $\mathbf{b}$ are perpendicular to each other (i.e. the angle between them is $\frac{\pi}{2}$ radians), then since $\cos\frac{\pi}{2} = 0$,

  $$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}|\,|\mathbf{b}| \cos\tfrac{\pi}{2} = 0.$$

- The dot product of a vector with itself gives the square of the magnitude of the vector, that is,

  $$\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|\,|\mathbf{a}| \cos 0 = |\mathbf{a}|^2.$$

The converse of the first property also holds: if $\mathbf{a}$ and $\mathbf{b}$ are two *non-zero* vectors such that $\mathbf{a} \cdot \mathbf{b} = 0$, then the definition of the dot product tells us that $\cos\theta = 0$; therefore $\theta = \frac{\pi}{2}$ and the vectors are perpendicular.

In a product of real numbers, $xy = 0$ implies that either x or y (or both) is zero. In contrast, for the dot product, $\mathbf{a} \cdot \mathbf{b} = 0$ gives an extra possibility: either $\mathbf{a}$ or $\mathbf{b}$ (or both) is the zero vector, or the angle between $\mathbf{a}$ and $\mathbf{b}$ is $\frac{\pi}{2}$ radians.

Properties of the dot product

The following are some important properties of the dot product of two vectors. They include the rules for manipulating dot products in algebraic expressions.

Properties of the dot product

Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be vectors, and let m be a scalar.

- $\mathbf{a} \cdot \mathbf{b}$ is a scalar.
- $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$, i.e. the dot product is commutative.
- $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ and $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$, i.e. the dot product is distributive over vector addition.
- $(m\mathbf{a}) \cdot \mathbf{b} = m(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (m\mathbf{b})$, i.e. a scalar can be 'moved through' a dot product.
- If neither $\mathbf{a}$ nor $\mathbf{b}$ is the zero vector, then $\mathbf{a} \cdot \mathbf{b} = 0$ if and only if $\mathbf{a}$ is perpendicular to $\mathbf{b}$.
- $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$.

These properties can all be derived from the definition of the dot product, but the derivations are not given here.

The following example shows how these properties can be used to simplify expressions.

Example 3

Expand the expression $\mathbf{x} \cdot \mathbf{y}$, given that $\mathbf{x} = 2\mathbf{u} + \mathbf{v}$ and $\mathbf{y} = \mathbf{u} - 5\mathbf{v}$. Calculate its value when $\mathbf{u}$ and $\mathbf{v}$ are perpendicular unit vectors.

Solution

$$\begin{aligned}\mathbf{x} \cdot \mathbf{y} &= (2\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} - 5\mathbf{v}) \\ &= 2(\mathbf{u} \cdot \mathbf{u}) - 10(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{u} - 5(\mathbf{v} \cdot \mathbf{v}) \\ &= 2(\mathbf{u} \cdot \mathbf{u}) - 9(\mathbf{u} \cdot \mathbf{v}) - 5(\mathbf{v} \cdot \mathbf{v}).\end{aligned}$$

Now $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 = 1$ and $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2 = 1$ when $\mathbf{u}$ and $\mathbf{v}$ are unit vectors. Furthermore, $\mathbf{u} \cdot \mathbf{v} = 0$ when $\mathbf{u}$ and $\mathbf{v}$ are perpendicular vectors. So when $\mathbf{u}$ and $\mathbf{v}$ are perpendicular unit vectors, we have

$$\mathbf{x} \cdot \mathbf{y} = 2 - 0 - 5 = -3.$$

Exercise 15

(a) Expand the expression $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b})$.

(b) Expand the expression $|\mathbf{a} + \mathbf{b}|^2$.

Recall that $|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}$.

Exercise 16

Suppose that $\mathbf{a}$ and $\mathbf{b}$ are perpendicular unit vectors.

(a) Find the value of m such that the two vectors $2\mathbf{a} + 3\mathbf{b}$ and $m\mathbf{a} + \mathbf{b}$ are perpendicular.

(b) Find the value of $|\mathbf{c}|$ if $\mathbf{c} = 3\mathbf{a} + 5\mathbf{b}$.

A word of caution: $(\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ is not in general the same as $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$. The vector $(\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ is a scaling of $\mathbf{c}$ by the number $\mathbf{a} \cdot \mathbf{b}$, whereas $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$ is a scaling of $\mathbf{a}$ by the number $\mathbf{b} \cdot \mathbf{c}$. Clearly these two vectors are not generally even parallel, let alone equal. For example, if $\mathbf{a} = \mathbf{b} = \mathbf{i}$ and $\mathbf{c} = \mathbf{j}$, then

$$(\mathbf{a} \cdot \mathbf{b})\mathbf{c} = (\mathbf{i} \cdot \mathbf{i})\mathbf{j} = \mathbf{j} \quad \text{but} \quad \mathbf{a}(\mathbf{b} \cdot \mathbf{c}) = \mathbf{i}(\mathbf{i} \cdot \mathbf{j}) = \mathbf{0}.$$

In general, if m is a scalar and $\mathbf{a}$ is a vector, we can write $m\mathbf{a}$ or $\mathbf{a}m$ as is convenient, although $m\mathbf{a}$ is more usual. Thus $\mathbf{a}(\mathbf{b} \cdot \mathbf{c})$ means the same as $(\mathbf{b} \cdot \mathbf{c})\mathbf{a}$.

Component form of the dot product

We saw in Section 2 that an arbitrary vector $\mathbf{a}$ in three dimensions may be expressed in terms of the Cartesian unit vectors as

$$\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}.$$

With this representation, vector addition and scaling become simple algebraic operations without any reference to diagrams. The definition of the dot product was expressed in terms of the magnitudes of two vectors and the angle between them. We will now see how to express the dot product in terms of components of vectors.

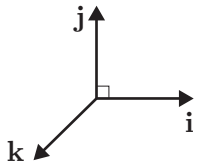


Figure 24 The Cartesian unit vectors (**k** points out of the plane of the page)

First, observe that by definition, **i**, **j** and **k** are unit vectors and are perpendicular to one another (see Figure 24). Thus

$$\mathbf{i}\cdot\mathbf{j}=\mathbf{j}\cdot\mathbf{i}=0,\quad \mathbf{i}\cdot\mathbf{k}=\mathbf{k}\cdot\mathbf{i}=0,\quad \mathbf{j}\cdot\mathbf{k}=\mathbf{k}\cdot\mathbf{j}=0,$$
$$\mathbf{i}\cdot\mathbf{i}=1,\quad \mathbf{j}\cdot\mathbf{j}=1,\quad \mathbf{k}\cdot\mathbf{k}=1.$$

If two vectors **a** and **b** have component forms $\mathbf{a}=a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k}$ and $\mathbf{b}=b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k}$, then the dot product of **a** and **b** may be written as

$$(a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k})\cdot(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k}).$$

We can now apply properties of the dot product and the above rules for combining **i**, **j** and **k** to this expression to obtain a very simple formula for the dot product of vectors in component form. Specifically, we have

$$\begin{aligned}&(a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k})\cdot(b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k})\\&\quad=a_1b_1(\mathbf{i}\cdot\mathbf{i})+a_1b_2(\mathbf{i}\cdot\mathbf{j})+a_1b_3(\mathbf{i}\cdot\mathbf{k})\\&\qquad+a_2b_1(\mathbf{j}\cdot\mathbf{i})+a_2b_2(\mathbf{j}\cdot\mathbf{j})+a_2b_3(\mathbf{j}\cdot\mathbf{k})\\&\qquad+a_3b_1(\mathbf{k}\cdot\mathbf{i})+a_3b_2(\mathbf{k}\cdot\mathbf{j})+a_3b_3(\mathbf{k}\cdot\mathbf{k})\\&\quad=a_1b_1+a_2b_2+a_3b_3.\end{aligned}$$

This extremely important formula is worth remembering.

Component form of the dot product

If $\mathbf{a}=a_1\,\mathbf{i}+a_2\,\mathbf{j}+a_3\,\mathbf{k}$ and $\mathbf{b}=b_1\,\mathbf{i}+b_2\,\mathbf{j}+b_3\,\mathbf{k}$, then the dot product of **a** and **b** is a scalar, given by

$$\mathbf{a}\cdot\mathbf{b}=a_1b_1+a_2b_2+a_3b_3. \tag{10}$$

In terms of vectors,

$$\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\cdot\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}=a_1b_1+a_2b_2+a_3b_3. \tag{11}$$

Exercise 17

If $\mathbf{a}=4\mathbf{i}+\mathbf{j}-5\mathbf{k}$ and $\mathbf{b}=\mathbf{i}-3\mathbf{j}+\mathbf{k}$, show that $\mathbf{a}\cdot\mathbf{b}=-4$. What does the negative sign tell us?

Angle between two vectors

The component form of the dot product has an important application in calculating the angle between two vectors. You have already seen that if $\mathbf{a}\cdot\mathbf{b}=0$ and neither **a** nor **b** is zero, then **a** and **b** are perpendicular. For instance, if $\mathbf{a}=2\mathbf{i}-\mathbf{j}+\mathbf{k}$ and $\mathbf{b}=2\mathbf{i}+3\mathbf{j}-\mathbf{k}$, then $\mathbf{a}\cdot\mathbf{b}=(2\times2)+(-1\times3)+(1\times-1)=0$, so the angle between **a** and **b** is $\pi/2$ radians. In general, the equation defining the dot product of **a** and **b**, that is, $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\,|\mathbf{b}|\cos\theta$, gives the following simple expression for finding the angle between **a** and **b**.

Angle between two vectors

The angle θ between any two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}}, \tag{12}$$

where $0 \leq \theta \leq \pi$.

Example 4

Let $\mathbf{a} = \mathbf{i} + \mathbf{j} + \sqrt{2}\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$ and $\mathbf{c} = \mathbf{i} + \mathbf{j} - \sqrt{2}\mathbf{k}$.

(a) Find the angle between the vector $\mathbf{a}$ and the x-axis.

(b) Find the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$.

(c) Show that $\mathbf{c}$ is perpendicular to $\mathbf{a}$.

Solution

(a) The direction of the x-axis is the same as the direction of $\mathbf{i}$, and the angle θ between $\mathbf{a}$ and $\mathbf{i}$ is given by

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{i}}{|\mathbf{a}|\,|\mathbf{i}|} = \frac{a_1}{|\mathbf{a}|} = \frac{1}{\sqrt{1+1+2}} = \frac{1}{2}.$$

Thus the angle between $\mathbf{a}$ and the x-axis is $\pi/3$ radians.

(b) We have $|\mathbf{a}| = \sqrt{1+1+2} = 2$, $|\mathbf{b}| = \sqrt{4+1+1} = \sqrt{6}$ and

$$\mathbf{a}\cdot\mathbf{b} = (1\times 2) + (1\times 1) + (\sqrt{2}\times 1) = 3 + \sqrt{2}.$$

Therefore the angle θ between $\mathbf{a}$ and $\mathbf{b}$ is given by

$$\cos\theta = \frac{3+\sqrt{2}}{2\times\sqrt{6}} = 0.901\,05,$$

so $\theta = 0.4486$ (radians).

(c) To test whether $\mathbf{a}$ and $\mathbf{c}$ are perpendicular, we calculate their dot product:

$$\begin{aligned}\mathbf{a}\cdot\mathbf{c} &= (\mathbf{i}+\mathbf{j}+\sqrt{2}\mathbf{k})\cdot(\mathbf{i}+\mathbf{j}-\sqrt{2}\mathbf{k})\\ &= (1\times 1) + (1\times 1) + (\sqrt{2}\times\sqrt{2})\\ &= 0.\end{aligned}$$

Since $\mathbf{a}\cdot\mathbf{c} = 0$ and $\mathbf{a}$ and $\mathbf{c}$ are non-zero vectors, $\mathbf{c}$ is perpendicular to $\mathbf{a}$.

Exercise 18

Consider the vectors $\mathbf{a} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ and $\mathbf{b} = -\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$.

Find the magnitudes of $\mathbf{a}$ and $\mathbf{b}$, and the angle between them.

Resolving a vector into components

The dot product has a useful geometric interpretation.

Exercise 19

If $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$, find the values of $\mathbf{a}\cdot\mathbf{i}$, $\mathbf{a}\cdot\mathbf{j}$ and $\mathbf{a}\cdot\mathbf{k}$.

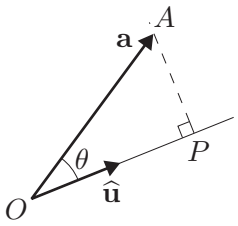


Figure 25 Finding the component of a vector in an arbitrary direction

Note that $\mathbf{a}\cdot\widehat{\mathbf{u}}$ will be negative if $\theta > \frac{\pi}{2}$, i.e. if P and $\widehat{\mathbf{u}}$ lie on opposite sides of O.

The solution to Exercise 19 shows the important fact that the **i**-component of any vector **a** may be found by taking the dot product $\mathbf{a}\cdot\mathbf{i}$. The **j**- and **k**-components can be found similarly (by taking dot products with **j** and **k**, respectively).

We can also find the components of a vector in other directions. Suppose that a vector **a**, represented by $\overrightarrow{OA}$, makes an angle θ with a *unit vector* $\widehat{\mathbf{u}}$ (see Figure 25). Draw the line AP perpendicular to the direction of $\widehat{\mathbf{u}}$. Then the distance OP is seen from simple trigonometry to be $|\mathbf{a}|\cos\theta$. Now observe that the dot product of **a** and $\widehat{\mathbf{u}}$ is

$$\mathbf{a}\cdot\widehat{\mathbf{u}} = |\mathbf{a}|\,|\widehat{\mathbf{u}}|\cos\theta = \pm|OP| \quad (\text{since } |\widehat{\mathbf{u}}| = 1).$$

The signed distance $\pm|OP|$ represents the *component* of **a** in the direction of $\widehat{\mathbf{u}}$.

The **component** of a vector **a** in the direction of an arbitrary *unit vector* $\widehat{\mathbf{u}}$ is $\mathbf{a}\cdot\widehat{\mathbf{u}}$.

Exercise 20

Consider the vectors $\mathbf{a} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ and $\mathbf{b} = -\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$.

(a) Which of the vectors

$$\mathbf{c} = -\mathbf{i} + \mathbf{j} + 3\mathbf{k}, \quad \mathbf{d} = -2\mathbf{i} + \mathbf{k}, \quad \mathbf{e} = -\mathbf{i} - \mathbf{j} - \mathbf{k},$$

is perpendicular to **a**?

(b) Find the component of the vector $\mathbf{a} + 2\mathbf{b}$ in the direction of the line joining the origin to the point $(1, 1, 1)$.

Resolving vectors will be a vital technique in this and subsequent units, and sometimes you will need to be able to resolve a vector into components in directions other than horizontal and vertical.

The dot product method of obtaining components always works, but a geometric view is also useful. This follows because the component of a vector **a** in the direction of a unit vector $\widehat{\mathbf{u}}$ is

$$\mathbf{a}\cdot\widehat{\mathbf{u}} = |\mathbf{a}|\cos\theta,$$

where θ is the angle between **a** and $\widehat{\mathbf{u}}$. We summarise the method as a procedure.

Procedure 1 Resolving a vector into components

Given a vector $\mathbf{a}$ and a unit vector $\hat{\mathbf{u}}$, to find the component of $\mathbf{a}$ in the direction of $\hat{\mathbf{u}}$, carry out the following steps.

1. Find (usually from a diagram) the angle θ between $\mathbf{a}$ and $\hat{\mathbf{u}}$ (with $0 \leq \theta \leq \pi$).
2. The component of the vector $\mathbf{a}$ in the direction of the unit vector $\hat{\mathbf{u}}$ is $|\mathbf{a}| \cos\theta$.
3. If necessary (e.g. if $\theta > \pi/2$), use trigonometric formulas from the Handbook to simplify the result.

Example 5

Figure 26 shows a line inclined at an angle α to the x-axis, and unit vectors $\mathbf{i}$ and $\mathbf{j}$ aligned along and perpendicular to the line. The vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ have magnitudes 1, 1.5, 1.5 and 2, respectively.

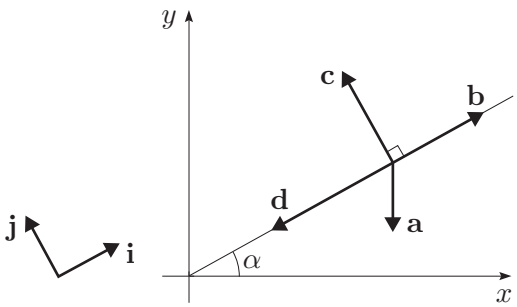


Figure 26

Resolve each of the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ into their $\mathbf{i}$- and $\mathbf{j}$-components.

Solution

We must find the angles between $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ and the unit vectors $\mathbf{i}$ and $\mathbf{j}$. First notice that $\mathbf{b}$ points in the direction $\mathbf{i}$, so $\mathbf{b} = 1.5\mathbf{i}$. Similarly, $\mathbf{d}$ points in the direction $-\mathbf{i}$, so $\mathbf{d} = -2\mathbf{i}$. Also, $\mathbf{c}$ points in the direction $\mathbf{j}$, so $\mathbf{c} = 1.5\mathbf{j}$.

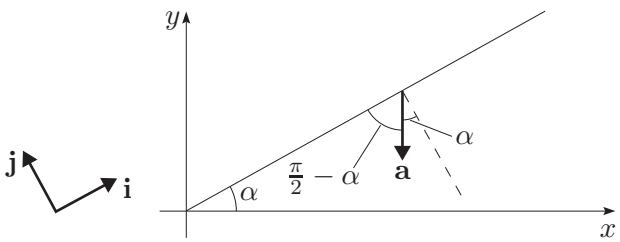


Figure 27

The remaining vector, $\mathbf{a}$, makes an angle $\frac{\pi}{2} - \alpha$ with $-\mathbf{i}$, and an angle α with $-\mathbf{j}$ (see Figure 27). Hence the $\mathbf{i}$-component of $\mathbf{a}$ is

$$-|\mathbf{a}| \cos\left(\tfrac{\pi}{2} - \alpha\right) = -\sin\alpha,$$

and the $\mathbf{j}$-component of $\mathbf{a}$ is

$$-|\mathbf{a}| \cos\alpha = -\cos\alpha.$$

Therefore

$$\mathbf{a} = -\sin\alpha\,\mathbf{i} - \cos\alpha\,\mathbf{j}.$$

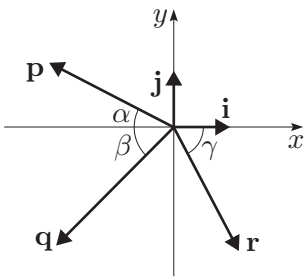


Figure 28

Exercise 21

The vectors $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ in Figure 28 have magnitudes 2.5, 3 and 2.5, respectively. Resolve $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ into their $\mathbf{i}$- and $\mathbf{j}$-components.

2.3 The cross product

You have seen that the dot product of two vectors is a scalar (i.e. a real number). In contrast, the *cross product* of two vectors is a *vector*, whose direction is perpendicular to both. The cross product has numerous applications in geometry and mechanics, as you will see later in the module.

The product $\mathbf{a} \times \mathbf{b}$ is read as 'a cross b'.

The **cross product** of two vectors $\mathbf{a}$ and $\mathbf{b}$ is

$$\mathbf{a} \times \mathbf{b} = (|\mathbf{a}|\,|\mathbf{b}| \sin\theta)\,\widehat{\mathbf{n}},$$

where θ $(0 \leq \theta \leq \pi)$ is the angle between the directions of $\mathbf{a}$ and $\mathbf{b}$, and $\widehat{\mathbf{n}}$ is a unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$, whose sense is given by the right-hand grip rule as shown in Figure 29(a).

The angle θ between two vectors $\mathbf{a}$ and $\mathbf{b}$ lies in the range $0 \leq \theta \leq \pi$, so $\sin\theta \geq 0$ and hence $|\mathbf{a}|\,|\mathbf{b}| \sin\theta \geq 0$. So the cross product of $\mathbf{a}$ and $\mathbf{b}$ is a vector with magnitude $|\mathbf{a}|\,|\mathbf{b}| \sin\theta$ and direction defined by $\widehat{\mathbf{n}}$. The direction of $\widehat{\mathbf{n}}$ is the direction in which the fingers in the grip in Figure 29(a) would advance when turned from $\mathbf{a}$ towards $\mathbf{b}$ through the angle θ. Notice that $\widehat{\mathbf{n}}$ is not defined if $\mathbf{a}$ and $\mathbf{b}$ are parallel or if $\mathbf{a}$ or $\mathbf{b}$ is the zero vector; but in these cases $|\mathbf{a}|\,|\mathbf{b}| \sin\theta = 0$, so we take $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. The cross product is also called the **vector product**, which stresses the fact that $\mathbf{a} \times \mathbf{b}$ is a vector.

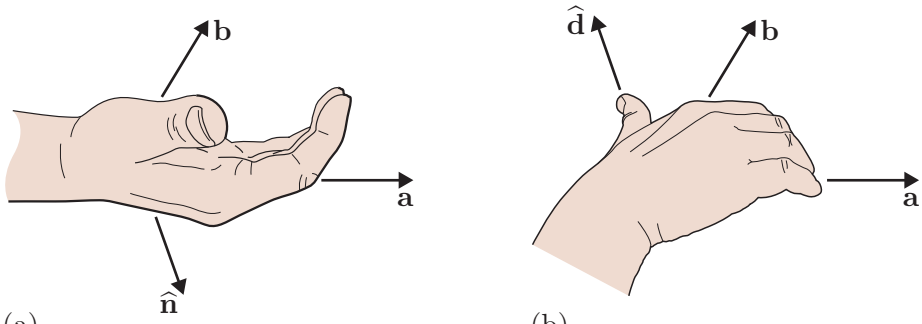


Figure 29 Using the right-hand grip rule for the cross product: (a) $\mathbf{a} \times \mathbf{b}$; (b) $\mathbf{b} \times \mathbf{a}$

The order of writing down $\mathbf{a}$ and $\mathbf{b}$ is very important. According to the right-hand grip rule, $\mathbf{b} \times \mathbf{a}$ is a vector in the direction opposite to $\mathbf{a} \times \mathbf{b}$ (see Figure 29(b)):

$$\mathbf{b} \times \mathbf{a} = (|\mathbf{b}|\,|\mathbf{a}| \sin\theta)\,\widehat{\mathbf{d}} = -(|\mathbf{b}|\,|\mathbf{a}| \sin\theta)\,\widehat{\mathbf{n}} = -(\mathbf{a} \times \mathbf{b}).$$

Exercise 22

Three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ lie in the (x, y)-plane. Their magnitudes are 2, 3 and 4 units, respectively, their directions make angles $\frac{\pi}{6}$, $\frac{\pi}{3}$ and $\frac{\pi}{6}$ radians, respectively, with the positive x-axis, and they have positive $\mathbf{j}$-components. Use the definition of the cross product to find the vectors $\mathbf{u} \times \mathbf{v}$, $\mathbf{u} \times \mathbf{w}$ and $\mathbf{v} \times \mathbf{w}$.

Exercise 22 illustrates an important property of the cross product. We know that if two vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel, then the angle θ between their directions is zero or π radians, so the cross product of $\mathbf{a}$ and $\mathbf{b}$ is the zero vector, because the magnitude of the vector, that is, $|\mathbf{a}|\,|\mathbf{b}| \sin\theta$, is zero. The converse also holds: if $\mathbf{a}$ and $\mathbf{b}$ are two non-zero vectors such that $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, then the definition of the cross product tells us that $\sin\theta = 0$; therefore $\theta = 0$ or $\theta = \pi$, and the vectors are parallel. We can also deduce that

$$\mathbf{a} \times \mathbf{a} = \mathbf{0} \quad \text{for any vector } \mathbf{a}.$$

So we can test for perpendicular vectors by using the dot product and for parallel vectors by using the cross product.

Properties of the cross product

The following are some important properties of the cross product of two vectors. They include the rules for manipulating cross products in algebraic expressions.

Properties of the cross product

Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be vectors, and let m be a scalar.

- $\mathbf{a} \times \mathbf{b}$ is a vector.
- $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$, i.e. the cross product is *not* commutative – the order does matter.
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c})$ and $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c})$, i.e. the cross product is distributive over vector addition.
- $(m\mathbf{a}) \times \mathbf{b} = m(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (m\mathbf{b})$, i.e. a scalar can be 'moved through' a cross product.
- If neither $\mathbf{a}$ nor $\mathbf{b}$ is the zero vector, then $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a}$ and $\mathbf{b}$ are parallel.
- $\mathbf{a} \times \mathbf{a} = \mathbf{0}$.
- In general, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

These properties can all be derived from the definition of the cross product, but the derivations are not given here.

Component form of the cross product

If two vectors $\mathbf{a}$ and $\mathbf{b}$ have component forms $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ and $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$, what is the component form of $\mathbf{a} \times \mathbf{b}$? We begin with an exercise.

Exercise 23

(a) Show that $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

(b) Calculate $\mathbf{j} \times \mathbf{i}$, $\mathbf{k} \times \mathbf{j}$ and $\mathbf{i} \times \mathbf{k}$.

(c) Calculate $\mathbf{i} \times \mathbf{i}$, $\mathbf{j} \times \mathbf{j}$ and $\mathbf{k} \times \mathbf{k}$.

(d) Expand and simplify

$$(\mathbf{i} + \mathbf{k}) \times (\mathbf{i} + \mathbf{j} + \mathbf{k}) \quad \text{and} \quad (\mathbf{i} \times (\mathbf{i} + \mathbf{k})) - ((\mathbf{i} + \mathbf{j}) \times \mathbf{k}).$$

(e) For two non-zero non-parallel vectors $\mathbf{a}$ and $\mathbf{b}$, simplify

$$(\mathbf{a} + \mathbf{b}) \times (\mathbf{a} + 2\mathbf{b}) \quad \text{and} \quad (\mathbf{a} + \mathbf{b}) \times (\mathbf{a} - \mathbf{b}).$$

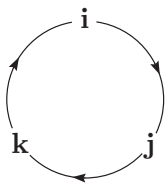


Figure 30 The cyclic pattern for the cross product

The cyclic pattern of the products $\mathbf{i} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{k}$, $\mathbf{k} \times \mathbf{i}$ and of the products $\mathbf{i} \times \mathbf{k}$, $\mathbf{k} \times \mathbf{j}$, $\mathbf{j} \times \mathbf{i}$, as demonstrated in Exercise 23, can be remembered using Figure 30. For example, if we go round the circle clockwise starting at $\mathbf{i}$, we have

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \quad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$$

However, if we go in an anticlockwise direction, the cross products are negative:

$$\mathbf{i} \times \mathbf{k} = -\mathbf{j}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \mathbf{j} \times \mathbf{i} = -\mathbf{k}.$$

Given $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ and $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$, the cross product $\mathbf{a} \times \mathbf{b}$ may be written as

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= (a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}) \times (b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}) \\ &= (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}.\end{aligned}$$

We highlight this important formula.

Component form of the cross product

If $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ and $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$, then

$$\mathbf{a} \times \mathbf{b} = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}. \qquad (13)$$

Alternatively,

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \qquad (14)$$

This formula is not easy to remember or use in this form. Another quick way to evaluate cross products is to use determinants. This method will be introduced in Unit 4 when we discuss determinants. If you already know this method, then we suggest that you continue to use it.

The determinant form of the cross-product $\mathbf{a} \times \mathbf{b}$, where $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ and $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$, can be written as

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \\ &= (a_2b_3 - a_3b_2)\mathbf{i} \\ &\quad + (a_3b_1 - a_1b_3)\mathbf{j} \\ &\quad + (a_1b_2 - a_2b_1)\mathbf{k}.\end{aligned}$$

Exercise 24

If

$$\mathbf{a} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}, \quad \mathbf{b} = -\mathbf{i} + 2\mathbf{j} + 4\mathbf{k} \quad \text{and} \quad \mathbf{c} = -4\mathbf{i} + 6\mathbf{j} - 2\mathbf{k},$$

find $\mathbf{a} \times \mathbf{b}$, $\mathbf{a} \times \mathbf{c}$ and $\mathbf{b} \times \mathbf{c}$. From your results, what can you say about $\mathbf{a}$ and $\mathbf{c}$?

Exercise 25

If

$$\mathbf{a} = 2\mathbf{i} + 2\mathbf{j} + \mathbf{k} \quad \text{and} \quad \mathbf{b} = 4\mathbf{i} + 4\mathbf{j} - 7\mathbf{k},$$

find a unit vector whose direction is perpendicular to the directions of both $\mathbf{a}$ and $\mathbf{b}$.

Geometric applications

We close this subsection with some useful geometric applications of the cross product. The following example is the first step.

Example 6

Any two non-zero and non-parallel vectors $\mathbf{a}$ and $\mathbf{b}$ define a parallelogram, as shown in Figure 31. Express the area of the parallelogram in terms of $\mathbf{a} \times \mathbf{b}$.

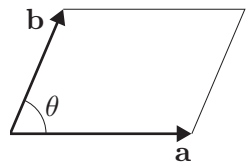


Figure 31 A parallelogram defined by two vectors

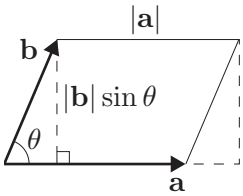


Figure 32 Finding the area of a parallelogram

Solution

The area A of the parallelogram defined by the two vectors $\mathbf{a}$ and $\mathbf{b}$ is the same as the area of the rectangle of height $|\mathbf{b}|\sin\theta$ and width $|\mathbf{a}|$ (see Figure 32). Thus $A = |\mathbf{a}|\,|\mathbf{b}|\sin\theta$, and this is the magnitude of $\mathbf{a}\times\mathbf{b}$. So

$$A = |\mathbf{a}\times\mathbf{b}|.$$

Area of a parallelogram

The area of a parallelogram with sides defined by vectors $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a}\times\mathbf{b}|$.

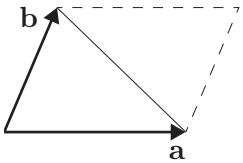


Figure 33 Finding the area of a triangle

This idea is easily extended for the area of a triangle. Any two non-zero non-parallel vectors $\mathbf{a}$ and $\mathbf{b}$ define a triangle (see Figure 33). The area of this triangle is half that of the corresponding parallelogram, so it is $\frac{1}{2}|\mathbf{a}\times\mathbf{b}|$.

Area of a triangle

The area of a triangle with sides defined by vectors $\mathbf{a}$ and $\mathbf{b}$ is $\frac{1}{2}|\mathbf{a}\times\mathbf{b}|$.

A *parallelepiped* is like a distorted brick. All of its faces are parallelograms.

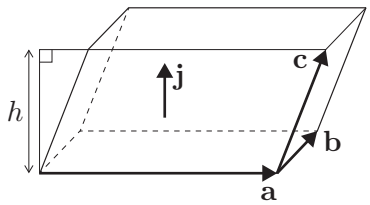


Figure 34 Finding the volume of a parallelepiped

Using the formula for the area of a parallelogram, we can find the volume of a parallelepiped (see Figure 34).

The base is a parallelogram (assumed to be in the (x,z)-plane) defined by the vectors $\mathbf{a}$ and $\mathbf{b}$. The base therefore has an area equal to the magnitude of $\mathbf{a}\times\mathbf{b}$. Now the vertical height h is the component of the vector $\mathbf{c}$ in the direction of the Cartesian unit vector $\mathbf{j}$ pointing vertically upwards, that is, it is the y-component of $\mathbf{c}$, given by $\mathbf{c}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{c}$. So the volume of the parallelepiped is $|\mathbf{a}\times\mathbf{b}|(\mathbf{j}\cdot\mathbf{c})$. But the vector product $\mathbf{a}\times\mathbf{b}$ points vertically upwards and can therefore be expressed as $|\mathbf{a}\times\mathbf{b}|\,\mathbf{j}$. Hence we have

$$|\mathbf{a}\times\mathbf{b}|(\mathbf{j}\cdot\mathbf{c}) = (|\mathbf{a}\times\mathbf{b}|\,\mathbf{j})\cdot\mathbf{c} = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}.$$

Of course, the scalar $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ can be negative if one of the defining vectors $\mathbf{a}$ or $\mathbf{b}$ is chosen to be in the direction opposite to the one chosen in Figure 34, or if the order of the cross product is reversed. We use modulus signs to ensure that the volume comes out positive:

$$\begin{aligned}\text{volume of parallelepiped} &= \text{base area} \times \text{vertical height } h \\ &= |(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}|. \qquad (15)\end{aligned}$$

The scalar quantity $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ is an example of a *scalar triple product*.

Exercise 26

Suppose that the vectors $\mathbf{r}$ and $\mathbf{s}$ are directed towards north and north-east, respectively, and define $\mathbf{r} \times \mathbf{s} = \mathbf{t}$.

(a) What is the direction of $\mathbf{t}$?

(b) In what direction is $\mathbf{s} \times \mathbf{r}$?

(c) In what direction is $\mathbf{t} \times \mathbf{r}$?

(d) If $|\mathbf{r}| = |\mathbf{s}| = 1$, what is $|\mathbf{t}|$?

(e) Calculate the vector $\mathbf{t} \times (\mathbf{r} \times \mathbf{s})$.

(f) If $|\mathbf{r}| = |\mathbf{s}| = 1$, what is the value of $\mathbf{r} \cdot \mathbf{s}$?

(g) If $|\mathbf{r}| = |\mathbf{s}| = 1$, what is the value of $\mathbf{s} \cdot (\mathbf{t} \times \mathbf{r})$?

Exercise 27

Find the area of the triangle ABC, where the coordinates of A, B and C are $(2, 1, -3)$, $(1, 0, 2)$ and $(4, -2, -1)$.

3 Modelling forces

This section shows how four common types of force can be modelled: the force of gravity, the force exerted by a surface on an object in contact with it, the tension force due to a string, and the friction force between two surfaces. These forces and the situations in which they occur are modelled and analysed in Subsections 3.2–3.5. In Subsection 3.6 we look at the forces on two particle systems. First, however, we look at one way of modelling the objects on which forces act.

3.1 Particles

When we create a mathematical model, the aim is to simplify the real situation being modelled so that only the essential features are included. This enables us to analyse the situation mathematically. In mechanics, the most important things to model are the forces acting on objects, and throughout this unit and the other mechanics units you will see how to do this. However, we also need to model the objects on which the forces act. The simplest model for an object is a *particle*.

A **particle** is a material object whose size and internal structure may be neglected. It has mass but no size, thus occupies a single point in space. A particle is often represented in diagrams by a black dot.

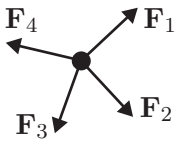


Figure 35 A force diagram

Note that in force diagrams, arrows are usually drawn with arbitrary lengths. This contrasts with the usual convention for vectors, where length indicates magnitude.

Observation has shown that each force acting on an object can be modelled as acting *at a particular point* on the object, this point being referred to as the **point of action** of the force. In situations where a particle model is appropriate, all the forces acting on the object are modelled as acting through the point in space occupied by the particle. It is conventional to show these forces in diagrams – known as **force diagrams** – by vector arrows whose tails coincide with the particle and whose directions correspond to the directions in which the forces act, as Figure 35 illustrates.

When several forces are acting on a particle, observation has shown that the overall effect of these forces can be represented by a single vector given by the sum of the vectors representing the individual forces. In this unit, we deal with objects that do not move, that is, objects in **equilibrium**. For a particle in equilibrium, the forces acting on it must balance each other (or else it would move), so we have the following important condition.

This condition was first stated by Isaac Newton as part of his first law of motion.

We often say that the sum of the forces is zero, with the implication that this means the zero vector.

Equilibrium condition for a particle

A particle subjected to forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ is in equilibrium if the forces sum to the zero vector, that is,

$$\sum_{i=1}^{n} \mathbf{F}_i = \mathbf{0}. \tag{16}$$

3.2 Weight

When you hold a shoe, your fingers experience a force. The shoe, like all objects, has a force associated with it, and if you do not provide opposition to this force in holding the shoe, it will fall to the ground. But what is the source of the force exerted by the shoe?

This force is due to the attraction of the shoe to the Earth. The force of attraction of objects to the Earth is called the **force of gravity** or the **gravitational force**. The gravitational force acting on a particular object is not constant, but depends on the position of the object relative to the Earth: there is a small variation of this force with height above ground (or depth below ground), and there is an even smaller variation with latitude and longitude. When applied to a particular object, this force is called the *weight* of the object. In this module, we assume that the weight of a particular object is constant near the Earth's surface.

In everyday speech, the words *mass* and *weight* are interchangeable. Mathematically, however, they are different. The **mass** of an object is the amount of matter in the object and is independent of the object's position in the Universe; it is a *scalar* quantity, measured in *kilograms* (kg) in the SI system. The **weight** of an object is the gravitational force on the object, and is dependent on where the object is situated; it is a *vector* quantity, whose magnitude is measured in *newtons* (N) in the SI system and whose direction is downwards towards the centre of the Earth.

Mass and weight are, however, related in that an object of mass m has weight of magnitude mg, where g is a constant known as the **magnitude of the acceleration due to gravity**. Near the Earth's surface, g has the value approximately $9.81\,\mathrm{m\,s^{-2}}$, and we assume this value for g throughout this module. If the Cartesian unit vector $\mathbf{j}$ points vertically upwards from the surface of the Earth, then the weight $\mathbf{W}$ of an object of mass m is $mg(-\mathbf{j})$ (where we need the negative sign because the force of gravity acts vertically downwards, that is, the weight acts vertically downwards).

The relationship between mass and weight is based on Newton's second law of motion, which is discussed in Unit 3.

An object of mass m has weight $\mathbf{W}$ of magnitude $|\mathbf{W}| = mg$, where g is the magnitude of the acceleration due to gravity, with direction towards the centre of the Earth. If the object is modelled as a particle, the force of gravity on the object can be illustrated by the force diagram in Figure 36.

Figure 36 A force diagram illustrating the weight of an object

Recall that a vector can be represented by its magnitude times a unit vector in the direction of the vector, that is, $\mathbf{a} = |\mathbf{a}|\,\widehat{\mathbf{a}}$.

Exercise 28

What is the weight (in newtons) of a particle of mass 3 kg in a coordinate system where the $\mathbf{k}$-direction is vertically downwards?

When modelling forces acting on objects, it is often convenient to define Cartesian unit vectors and to express the force vectors in component form, that is, to resolve the vectors into their components. These Cartesian unit vectors define the directions of the axes in a Cartesian coordinate system, so we often refer to the process of defining Cartesian unit vectors as **choosing axes**.

Exercise 29

Later in this unit we will find it convenient to use axes that are not horizontal and vertical. Express the weight $\mathbf{W}$ of a particle of mass 15 kg in terms of the Cartesian unit vectors $\mathbf{i}$ and $\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ both lie in a vertical plane and are oriented as shown in Figure 37.

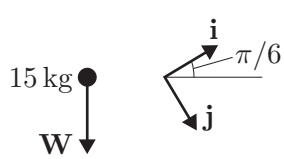


Figure 37 Using axes that are not horizontal and vertical

In Exercise 29, 'nice' angles (i.e. multiples of $\frac{\pi}{6}$ $(30°)$) were chosen in order to help you to evaluate the sine and cosine involved without having to use a calculator. The sines and cosines of some 'convenient' angles are as follows:

These values are given in the Handbook.

$$\sin 0 = 0, \quad \sin\tfrac{\pi}{6} = \tfrac{1}{2}, \quad \sin\tfrac{\pi}{4} = \tfrac{1}{\sqrt{2}}, \quad \sin\tfrac{\pi}{3} = \tfrac{\sqrt{3}}{2}, \quad \sin\tfrac{\pi}{2} = 1,$$
$$\cos 0 = 1, \quad \cos\tfrac{\pi}{6} = \tfrac{\sqrt{3}}{2}, \quad \cos\tfrac{\pi}{4} = \tfrac{1}{\sqrt{2}}, \quad \cos\tfrac{\pi}{3} = \tfrac{1}{2}, \quad \cos\tfrac{\pi}{2} = 0.$$

Sometimes, obtuse angles are used; sines and cosines of such angles can be derived from the addition formulas given in the Handbook. For example,

$$\cos\tfrac{2\pi}{3} = \cos\left(\tfrac{\pi}{2} + \tfrac{\pi}{6}\right) = \cos\tfrac{\pi}{2}\cos\tfrac{\pi}{6} - \sin\tfrac{\pi}{2}\sin\tfrac{\pi}{6} = -\tfrac{1}{2}.$$

3.3 Normal reaction

Consider an empty coffee mug resting on a table. Let us model the mug as a particle and consider the forces acting on it. We know that one force, the mug's weight, is acting on the mug. But since the mug is at rest (i.e. not moving), the equilibrium condition for particles tells us that some other force(s) must be acting on the mug (so that all the forces acting on the mug sum to zero). The only possible source for another force on the mug is the table. The force exerted by the table on the mug, and indeed exerted by any surface on an object in contact with it, is called the **normal reaction force** or simply the **normal reaction**.

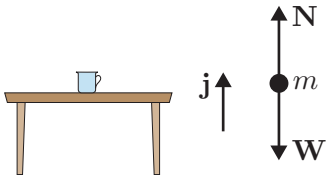


Figure 38 The forces on a mug resting on a table

The situation is illustrated in Figure 38, which shows not only the mug and table, but the corresponding force diagram (plus the Cartesian unit vector $\mathbf{j}$ pointing vertically upwards). The normal reaction force is denoted by $\mathbf{N}$, and the weight of the mug is denoted by $\mathbf{W}$. Using the equilibrium condition for particles, we have

$$\mathbf{W} + \mathbf{N} = \mathbf{0}.$$

If the mug has mass m, then $\mathbf{W} = mg\,(-\mathbf{j})$, hence

$$\mathbf{N} = -\mathbf{W} = mg\,\mathbf{j},$$

that is, the normal reaction is a force acting vertically upwards with the same magnitude as the weight of the mug.

The normal reaction force is remarkable in that it adjusts itself to the magnitude required. For example, if the coffee mug is replaced by a full pot of coffee, then the normal reaction increases (unless the weight of the coffee pot is too much for the table, in which case the table collapses and the pot is no longer at rest). Contrast this with the weight of an object, which is fixed and constant, regardless of what is happening to the object. Our basic modelling assumption is that the magnitude of the normal reaction force is potentially unlimited.

There is a normal reaction force whenever one object (e.g. a mug) presses on another (e.g. a table). Observation has shown that this force acts *normally* (i.e. at a right angle) to the surface at the point of contact between the objects. It therefore does not always act vertically upwards.

This explains the name *normal* reaction force.

For example, if the table on which the mug is resting is on an uneven floor, so that the table top makes an angle θ with the horizontal, then the normal reaction force makes an angle θ with the vertical, as shown in Figure 39. (In such a case there must be other forces acting on the mug if it is to remain in equilibrium. These other forces are discussed later.)

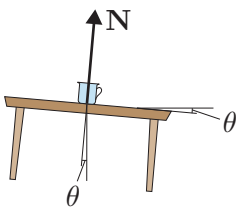


Figure 39 The normal reaction force acts at right angles to the surface

3.4 Tension

Consider a lamp hanging from a ceiling on an electric cable. Let us model the lamp as a particle. As in the case of the mug and table in the previous subsection, we know that there is a weight associated with the lamp, and that since the lamp is at rest, by the equilibrium condition some other force(s) must be acting on it. The only possible source for another force is

the cable, so the cable must exert a force on the lamp. The force exerted by the cable on the lamp is a vector quantity called the **tension force**.

Tension forces occur whenever objects are tautly joined, for example by cables, ropes, strings or threads. These cables and ropes can be modelled in different ways. For example, if we want to model the ceiling lamp and are interested only in the force in the cable, then we can model the cable as a **model string**, defined as an object possessing length, but no area, volume or mass, and which does not stretch (i.e. it is inextensible). On the other hand, if we are interested in how much the cable stretches under the weight of the lamp, then we can model the cable as a *model spring*, which has properties similar to those of a model string (i.e. it has no area, volume or mass), but allows extension. In this unit we consider only strings.

Springs are discussed in Unit 9.

The ceiling lamp example is illustrated in Figure 40. The tension force due to the model string is denoted by $\mathbf{T}$, and the weight of the lamp is denoted by $\mathbf{W}$. In a manner similar to the case of normal reaction forces, the equilibrium condition for particles gives

$$\mathbf{W} + \mathbf{T} = \mathbf{0}.$$

If the lamp has mass m, then $\mathbf{W} = mg\,(-\mathbf{j})$, hence

$$\mathbf{T} = -\mathbf{W} = mg\,\mathbf{j},$$

that is, tension is a force acting vertically upwards (along the length of the model string) with the same magnitude as the weight of the lamp.

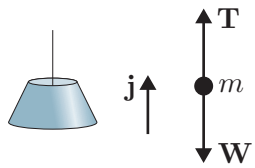


Figure 40 Modelling a lamp hanging from a ceiling

We assume that the tension force due to a model string acts along the length of the string and away from the point of its attachment to an object. As in the case of a normal reaction, the magnitude of this force (a scalar quantity, often referred to as the *tension in the string*) depends on the requirements necessary to maintain equilibrium, so it is potentially unlimited. (In reality, a string can exert only a certain tension force before it breaks, but a *model* string supports an unlimited tension force.)

A **model string** is an object with a fixed finite length, and no area, volume or mass, that exerts a force at the point of attachment.

The **tension force due to a string** is a vector quantity that acts along the length of the string away from the point of attachment.

As in the case of normal reaction forces, the tension force due to a string need not be vertically upwards, as the following example illustrates.

Example 7

A hanging flower basket of mass 4 kg is suspended by one cord from a porch and tied by another cord to the wall, as shown in Figure 41. Model the basket as a particle and the cords as model strings. What are the magnitudes of the tension forces due to the cords?

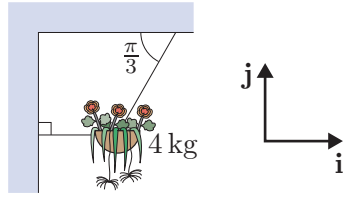


Figure 41 A flower basket hanging in a porch

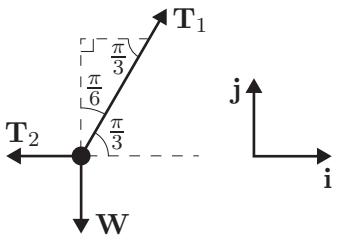


Figure 42 The force diagram for the hanging basket

Solution

We choose axes as shown in Figure 41. Note that we need choose only two axes because all the forces act in the same vertical plane. Denoting the tension forces by $\mathbf{T}_1$ and $\mathbf{T}_2$, and the weight of the basket by $\mathbf{W}$, we have the force diagram shown in Figure 42.

In the diagram, the angle between the vector $\mathbf{T}_1$ and the unit vector $\mathbf{j}$ is calculated by imagining the right-angled triangle shown, and using the fact that the angles of a triangle sum to π radians. The angle between the vectors $\mathbf{T}_1$ and $\mathbf{i}$ is $\frac{\pi}{3}$.

The equilibrium condition for particles tells us that

$$\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{W} = \mathbf{0}. \tag{17}$$

To progress further, we need to express the three forces in terms of the unit vectors $\mathbf{i}$ and $\mathbf{j}$. To do this, we apply Procedure 1, the technique for resolving vectors. Starting with the weight $\mathbf{W}$, where $|\mathbf{W}| = 4g$ and $\mathbf{W}$ acts in the direction $-\mathbf{j}$, we have

$$\mathbf{W} = -4g\,\mathbf{j}.$$

A vector $\mathbf{u} = \cos\alpha\,\mathbf{i} + \sin\alpha\,\mathbf{j}$ has magnitude

$$|\mathbf{u}| = \sqrt{\cos^2\alpha + \sin^2\alpha} = 1$$

and is thus a unit vector.

Similarly, the tension force $\mathbf{T}_1$ has magnitude $|\mathbf{T}_1|$ in the direction $\widehat{\mathbf{u}} = \cos\frac{\pi}{3}\,\mathbf{i} + \sin\frac{\pi}{3}\,\mathbf{j}$ and can be expressed as

$$\mathbf{T}_1 = |\mathbf{T}_1|\cos\tfrac{\pi}{3}\,\mathbf{i} + |\mathbf{T}_1|\sin\tfrac{\pi}{3}\,\mathbf{j} = \tfrac{1}{2}|\mathbf{T}_1|\,\mathbf{i} + \tfrac{\sqrt{3}}{2}|\mathbf{T}_1|\,\mathbf{j}.$$

Finally, the tension force $\mathbf{T}_2$ has magnitude $|\mathbf{T}_2|$ in the direction $-\mathbf{i}$ and can be written as

$$\mathbf{T}_2 = -|\mathbf{T}_2|\,\mathbf{i}.$$

Substituting in equation (17), the equilibrium condition is

$$\tfrac{1}{2}|\mathbf{T}_1|\,\mathbf{i} + \tfrac{\sqrt{3}}{2}|\mathbf{T}_1|\,\mathbf{j} - |\mathbf{T}_2|\,\mathbf{i} - 4g\,\mathbf{j} = \mathbf{0}.$$

Either by separating out the $\mathbf{i}$- and $\mathbf{j}$-components, or equivalently, taking the dot product with $\mathbf{i}$ and $\mathbf{j}$ in turn, gives the two scalar equations

$$\tfrac{1}{2}|\mathbf{T}_1| - |\mathbf{T}_2| = 0, \tag{18}$$

$$\tfrac{\sqrt{3}}{2}|\mathbf{T}_1| - 4g = 0. \tag{19}$$

Equation (19) gives

$$|\mathbf{T}_1| = 8g/\sqrt{3} \simeq 45.31.$$

Substituting this into equation (18) gives

$$|\mathbf{T}_2| = 4g/\sqrt{3} \simeq 22.66.$$

So the model predicts that the tension force due to the cord from the porch has magnitude about 45.3 N, and the tension force due to the cord from the wall has magnitude about 22.7 N, both correct to one decimal place.

The procedure that was used in Example 7 can be used to solve many problems in statics, and may be summarised as follows.

Procedure 2 Solving statics problems for particles

Given a statics problem, perform some or all of the following steps.

1. Draw a sketch of the physical situation, and annotate it with any relevant information. ◀ Draw picture ▶
2. Choose axes, and mark them on your sketch. ◀ Choose axes ▶
3. Draw a force diagram for each particle. ◀ Draw force diagram(s) ▶
4. Use the equilibrium condition on each particle and any other appropriate law(s) to obtain equation(s). ◀ Apply law(s) ▶
5. Solve the equation(s). ◀ Solve equation(s) ▶
6. Interpret the solution in terms of the original problem. ◀ Interpret solution ▶

In this unit, the steps in this procedure will often be identified (using the marginal abbreviations shown above) in the solutions to examples and exercises. However, the procedure is intended to be a guide rather than a rigid set of rules. For example, if it is not obvious which set of axes to choose, then draw the force diagram first, and the best choice may become more apparent. Try using the procedure in the following exercise.

Exercise 30

During December, a large plastic Christmas tree of mass 10 kg is suspended by its apex using two ropes attached to buildings either side of the high street of Treppendorf. The ropes make angles of $\pi/6$ and $\pi/4$ with the horizontal. Model the Christmas tree as a particle and the ropes as model strings. What are the magnitudes of the tension forces due to the two ropes?

3.5 Friction

Consider a book resting on a horizontal surface such as a table. There are two forces acting on the book: the weight downwards and the normal reaction upwards. Suppose that you push the book gently sideways (see Figure 43). If you do not push hard enough, the book will not move; it will remain in equilibrium. The force opposing motion along the surface is known as the **friction force**. It is considered to act parallel to the surface, that is, at right angles to the normal reaction, and in a direction that opposes any (possible) motion along that surface. Modelling the book as a particle, and denoting the pushing force by **P**, the friction force by **F**, the weight by **W** and the normal reaction by **N**, the force diagram for this example is shown in Figure 44.

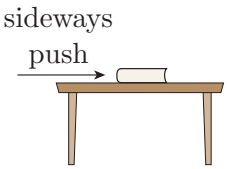


Figure 43 Pushing a book resting on a table

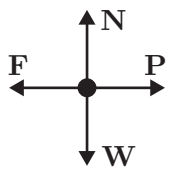
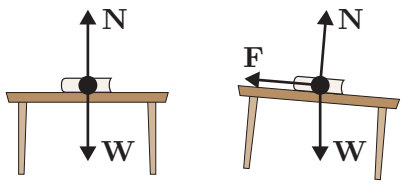


Figure 44 The force diagram for the book being pushed

Friction forces are caused by the roughness of even seemingly very smooth surfaces – a roughness that serves to inhibit the smooth movement of one surface over another. So friction forces are present only where there is movement or the possibility of movement. There is no friction force present when an object is resting on a horizontal surface, where the only two forces acting on the object are its weight and the normal reaction. But when an object is being pushed or pulled, or is resting on a sloping surface, then a friction force is present (see Figure 45).

In this unit we consider only cases where objects remain at rest, so that there is only the *possibility* of movement. Friction in cases where there *is* movement is considered in Unit 3.



no friction force | friction force present

Figure 45 The possible forces acting on the book

Unlike the normal reaction, which is potentially unlimited in magnitude, there is a limit to the magnitude of the friction force; if this limit is reached, then slipping occurs. The limiting value of the magnitude of the friction force depends almost entirely on the materials of the two surfaces and on the magnitude of the normal reaction force between them. It does not usually depend on the area of contact between the two surfaces, or on the angle at which the two surfaces are inclined to the horizontal. Experiments show that the limiting value of the magnitude of the friction force $\mathbf{F}$ (which *just* prevents slipping for two given surfaces) is approximately proportional to the magnitude of the normal reaction force $\mathbf{N}$ between the two surfaces. So on the verge of slipping, we have the scalar equation

$$|\mathbf{F}| = \mu|\mathbf{N}|,$$

where μ is the **coefficient of static friction**, which depends on the materials of the two surfaces. Some approximate values of μ for different materials are given in Table 1.

Table 1 Approximate coefficients of static friction

Surface	μ
Steel on steel (dry)	0.74
Steel on steel (oiled)	0.14
Plastic on plastic	0.35
Rubber on tarmac	0.85
Steel on wood	0.55
Wood on wood	0.42

Example 8

A steel fork of mass 0.05 kg rests on a horizontal wooden table. Model the fork as a particle. What is the maximum sideways force that can be applied before the fork starts to move?

Solution

The situation is illustrated in Figure 46. Since all the forces act in a vertical plane, we can choose axes as shown. The force diagram is also shown in the figure, where **F** is the friction force, **P** is the sideways force, **W** is the weight, and **N** is the normal reaction. The obvious choice for the directions of the unit vectors **i** and **j** is horizontal and vertical, respectively, as all the forces act in these directions.

◀ Draw picture ▶

◀ Choose axes ▶

◀ Draw force diagram(s) ▶

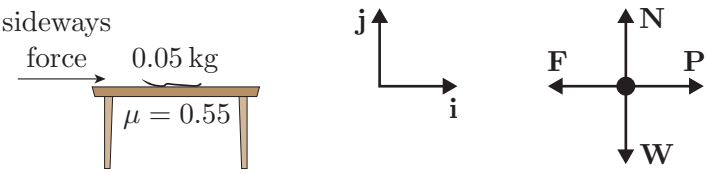


Figure 46 A force acting on a fork on a table, and the force diagram

The equilibrium condition for particles gives

◀ Apply law(s) ▶

$$\mathbf{F} + \mathbf{N} + \mathbf{P} + \mathbf{W} = \mathbf{0}. \tag{20}$$

To be able to use equation (20), we need to express the forces in terms of the unit vectors **i** and **j**. Looking at Figure 46, the forces can be written in component form as

◀ Solve equation(s) ▶

$$\mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}), \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{P} = |\mathbf{P}|\,\mathbf{i}, \quad \mathbf{W} = |\mathbf{W}|\,(-\mathbf{j}).$$

Resolving equation (20) in the **i**-direction gives

$$-|\mathbf{F}| + 0 + |\mathbf{P}| + 0 = 0,$$

so (as expected)

$$|\mathbf{F}| = |\mathbf{P}|.$$

Resolving equation (20) in the **j**-direction gives

$$|\mathbf{N}| = |\mathbf{W}|.$$

When the fork is on the point of moving (slipping), we have

$$|\mathbf{F}| = \mu|\mathbf{N}|,$$

where $\mu = 0.55$ is the coefficient of static friction for steel on wood.

Therefore when the fork is on the point of moving,

$$|\mathbf{P}| = |\mathbf{F}| = \mu|\mathbf{N}| = \mu|\mathbf{W}| = 0.55 \times 0.05g = 0.0275g \simeq 0.27.$$

So the model predicts that a sideways force of magnitude 0.27 N, correct to two decimal places, can be applied without moving the fork.

◀ Interpret solution ▶

Here is a summary of how we go about modelling problems that involve static friction, that is, problems involving friction but no motion.

Modelling static friction

Consider two surfaces in contact.

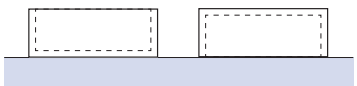
- The friction force $\mathbf{F}$ acts in a direction perpendicular to the normal reaction $\mathbf{N}$ between the surfaces and opposite to any possible motion along the common tangent to the surfaces.
- $|\mathbf{F}| \leq \mu|\mathbf{N}|$, where μ is a constant called the coefficient of static friction for the two surfaces involved.
- $|\mathbf{F}| = \mu|\mathbf{N}|$ when the object is on the verge of slipping. This equality is sometimes referred to as describing a situation of **limiting friction**.
- If one of the surfaces is designated as being **smooth**, it may be assumed that there is no friction present when this surface is in contact with another, regardless of the roughness of the other surface.

$|\mathbf{F}|$ cannot exceed its limiting value $\mu|\mathbf{N}|$. Slipping occurs if a friction force of magnitude greater than $\mu|\mathbf{N}|$ would be needed to prevent it.

Let us now apply these ideas to some examples, in which we will also apply the steps of Procedure 2. In most of the situations that we investigate, we will be concerned with limiting friction.

Exercise 31

A steel block of mass 0.5 kg rests on a horizontal dry steel surface (with coefficient of static friction $\mu = 0.74$) and is pulled by a horizontal force of 2 N. Model the block as a particle. Use Procedure 2 to determine whether the block will move. What is the magnitude of the friction force?



(a) (b)

Figure 47 A box without a lid: (a) with its base in contact with a table; (b) with its open top in contact with the table

Exercise 32

A shallow box made of a uniform material and without a lid can be placed on a horizontal table in two possible ways (as shown in Figure 47):

(a) with its base in contact with the table surface

(b) with its open top in contact with the table surface.

Which of these two positions requires the smaller sideways force to start the box slipping?

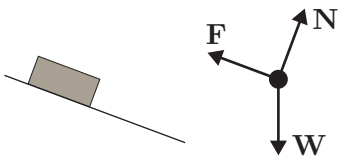


Figure 48 An object on an inclined plane, with its force diagram

Inclined planes

Consider now an object resting on a sloping plane surface, often referred to as an **inclined plane**, such as the one shown in Figure 48. Provided that the angle of inclination is not large, the object can remain at rest and does not slide down the slope. The forces acting on the object are its weight,

the normal reaction and friction. The weight **W** acts vertically downwards. The normal reaction **N** acts normally to the surface between the object and the slope. The friction force **F** is perpendicular to the normal reaction and hence parallel to the slope, and it acts up the slope to counteract the natural tendency of the object to move down the slope.

Example 9

A crate of empty bottles of total mass 30 kg is to be hauled by a rope up a ramp. The rope is parallel to the ramp, and the ramp makes an angle of $\pi/6$ radians with the horizontal. The coefficient of static friction between the plastic crate and the wooden ramp is 0.2.

What is the tension force due to the rope when the crate is on the point of moving upwards?

Solution

The situation is illustrated in Figure 49. ◀ Draw picture ▶

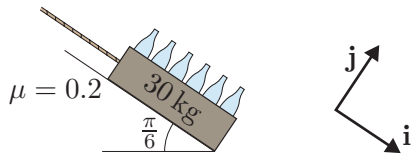


Figure 49 A crate of bottles being pulled up an inclined ramp

All the forces act in a vertical plane, so we need only two axes. We could choose **i** to be horizontal and **j** vertical as before, but it makes calculations easier if we choose **i** to be parallel to the slope and **j** perpendicular to it, as shown in Figure 49. This is because when we come to resolve the forces in the **i**- and **j**-directions, three of the four forces (all except **W**) will then act along one or other of the axes, making resolving them much simpler. ◀ Choose axes ▶

Modelling the crate as a particle and the rope as a model string, the force diagram is as shown in Figure 50, where **W** is the weight, **N** is the normal reaction, **F** is the friction force, and **T** is the tension force. ◀ Draw force diagram(s) ▶

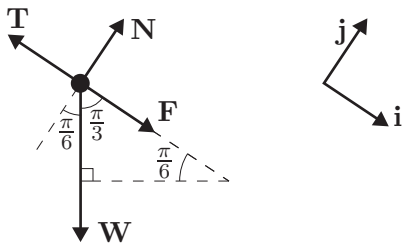


Figure 50 The force diagram for the crate

The equilibrium condition for particles gives ◀ Apply law(s) ▶

$$\mathbf{T} + \mathbf{N} + \mathbf{F} + \mathbf{W} = \mathbf{0}. \tag{21}$$

When the crate is on the point of moving, we have

$$|\mathbf{F}| = \mu|\mathbf{N}|,$$

where $\mu = 0.2$ is the coefficient of static friction.

◀ Solve equation(s) ▶

As before, the first step in solving the equations involves resolving the force vectors into components. In this case, three of the force vectors are aligned with the axes and can be written down immediately:

$$\mathbf{F} = |\mathbf{F}|\,\mathbf{i}, \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{T} = |\mathbf{T}|\,(-\mathbf{i}).$$

Note that $\mathbf{W} = |\mathbf{W}|\,\widehat{\mathbf{u}}$, where $\widehat{\mathbf{u}}$ is a unit vector. A unit vector in the plane is always of the form $\widehat{\mathbf{u}} = \cos\alpha\,\mathbf{i} \pm \sin\alpha\,\mathbf{j}$.

To find the weight of the crate in component form, we resolve:

$$\mathbf{W} = |\mathbf{W}|\cos\tfrac{\pi}{3}\,\mathbf{i} - |\mathbf{W}|\sin\tfrac{\pi}{3}\,\mathbf{j} = \tfrac{1}{2}|\mathbf{W}|\,\mathbf{i} - \tfrac{\sqrt{3}}{2}|\mathbf{W}|\,\mathbf{j}.$$

Now equation (21) can easily be resolved in the **i**-direction, giving

$$-|\mathbf{T}| + 0 + |\mathbf{F}| + \tfrac{1}{2}|\mathbf{W}| = 0,$$

so

$$|\mathbf{T}| = |\mathbf{F}| + \tfrac{1}{2}|\mathbf{W}|. \tag{22}$$

Resolving equation (21) in the **j**-direction gives

$$0 + |\mathbf{N}| + 0 - \tfrac{\sqrt{3}}{2}|\mathbf{W}| = 0,$$

so

$$|\mathbf{N}| = \tfrac{\sqrt{3}}{2}|\mathbf{W}|. \tag{23}$$

At the point of moving, $|\mathbf{F}| = 0.2|\mathbf{N}|$, and equations (22) and (23) give

$$|\mathbf{T}| = 0.2|\mathbf{N}| + \tfrac{1}{2}|\mathbf{W}| = 0.2 \times \tfrac{\sqrt{3}}{2}|\mathbf{W}| + \tfrac{1}{2}|\mathbf{W}|.$$

Thus, since $|\mathbf{W}| = 30g$,

$$|\mathbf{T}| = \left(\tfrac{\sqrt{3}}{10} + \tfrac{1}{2}\right) \times 30g \simeq 198.$$

◀ Interpret solution ▶

Therefore when the crate is on the point of moving, the model predicts that the tension force due to the rope is 198 N, to the nearest whole number, up the ramp.

Mathematically, different choices of axes make no difference to the final solution obtained to a mechanics problem. However, a sensible choice of axes, as in Example 9, can reduce the amount of calculation. You will find that with experience, you will be able to choose axes that reduce the work involved.

Choice of axes is discussed again in Unit 3.

Exercise 33

A full crate of bottles of mass 60 kg is at the top of the ramp described in Example 9, ready to be lowered down it. What force needs to be applied to the rope to keep the crate from sliding down the ramp?

Exercise 34

(a) A box of mass m is resting on a surface inclined at an angle α to the horizontal. If the box is on the point of slipping, what is the coefficient of static friction?

(b) Two identical mugs are placed on a tray. One mug is half full of coffee, while the other is empty. The tray is tilted slowly. Use your answer to part (a) to determine which mug will start to move first.

The result of Exercise 34(a) provides us with a technique for estimating the coefficient of static friction μ for two surfaces. Put the two surfaces in contact, and increase the angle of inclination from the horizontal until slipping begins. The tangent of the angle at which this happens is the required value of μ.

3.6 Two-particle systems

In the previous subsections we considered the action of forces on one particle and introduced the equilibrium condition for particles. In this subsection we extend these ideas to situations involving two or more particles. We will model such situations by considering the forces acting on each particle separately. Each particle is aware of only the forces acting on it, and is unaware of any forces acting on other particles.

We begin by introducing a new modelling device – the model pulley – then consider friction in the two-particle case.

Pulleys

The pulley is a common device with which you are probably familiar. You may have seen pulleys in use on building sites, for example, as an aid to raising or lowering heavy loads. The idea of a pulley is useful in modelling mechanics problems, as it enables us to model a change in direction of a tension force.

In diagrams, we use an idealised pulley as shown in Figure 51. In order to keep the model simple, we make simplifying assumptions, which are formally stated in the following definition.

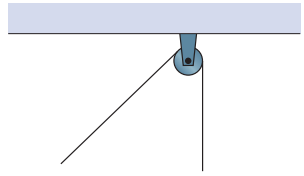


Figure 51 An idealised pulley

A **model pulley** is an object with no mass or size, over which a model string may pass without any resistance to motion. The magnitude of the tension in a string passing over a model pulley is the same either side of the pulley.

The point to remember is that the result of these assumptions implies that the tension forces due to the string on either side of the model pulley are equal in magnitude, that is, the magnitude of the tension in the string is the same on either side of the pulley (although the tension forces act in different directions).

A model pulley provides a reasonable model of an actual pulley, provided that its dimensions are small compared with the length of the rope or cable passing over it and that its weight is small compared with the other forces involved. Model pulleys can also be used to model a variety of situations that do not involve pulleys at all, but merely involve a change in direction of a tension force (such as when a rope is hanging over the edge of a building). Their use is illustrated by the following example and exercise.

Example 10

A sack of flour of mass 50 kg is lying on the floor of a mill, ready to be loaded into a cart. To help with the loading process, a light rope is attached to the sack, passes over a pulley fixed to the ceiling immediately above the sack, and is attached at its other end to a stone of mass 15 kg that hangs without touching the floor. The system is shown in Figure 52.

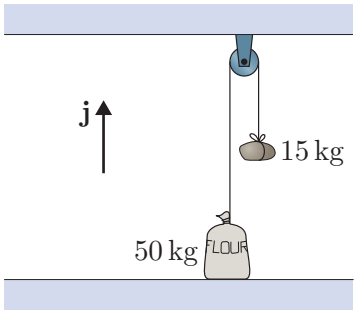


Figure 52 Raising a sack of flour using a pulley

Model the sack and the stone as particles, the pulley as a model pulley, and the rope as a model string, and consider the forces acting on each particle separately.

(a) Calculate the normal reaction of the floor on the sack.

(b) What force does the pulley exert on the ceiling?

Solution

◄ Choose axes ►

(a) All the forces are vertical, so we need only one axis, as shown in Figure 52.

◄ Draw force diagram(s) ►

The force diagrams for the sack and the stone are shown in Figure 53, where $\mathbf{W}_1$ and $\mathbf{W}_2$ represent the weights, $\mathbf{T}_1$ and $\mathbf{T}_2$ represent the tension forces, and $\mathbf{N}$ is the normal reaction of the floor on the sack.

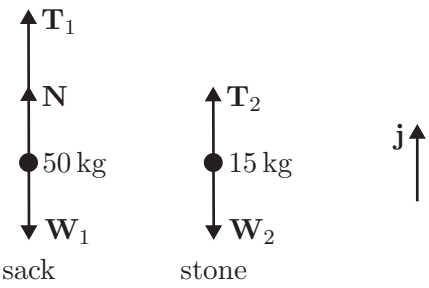


Figure 53 Force diagrams for the sack and the stone

◄ Apply law(s) ►

The equilibrium condition for particles gives

$$\mathbf{W}_1 + \mathbf{N} + \mathbf{T}_1 = \mathbf{0}, \tag{24}$$

$$\mathbf{W}_2 + \mathbf{T}_2 = \mathbf{0}. \tag{25}$$

Since the tension forces on either side of a model pulley have the same magnitude, we have

$$|\mathbf{T}_1| = |\mathbf{T}_2|.$$

◀ Solve equation(s) ▶

To solve the equations, the first step is to write the forces in component form as

$$\mathbf{W}_1 = |\mathbf{W}_1|\,(-\mathbf{j}), \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{T}_1 = |\mathbf{T}_1|\,\mathbf{j},$$
$$\mathbf{T}_2 = |\mathbf{T}_2|\,\mathbf{j}, \quad \mathbf{W}_2 = |\mathbf{W}_2|\,(-\mathbf{j}).$$

Then resolving equations (24) and (25) in the **j**-direction gives

$$-|\mathbf{W}_1| + |\mathbf{N}| + |\mathbf{T}_1| = 0,$$
$$-|\mathbf{W}_2| + |\mathbf{T}_2| = 0.$$

Therefore $|\mathbf{T}_1| = |\mathbf{T}_2| = |\mathbf{W}_2|$, so

$$|\mathbf{N}| = |\mathbf{W}_1| - |\mathbf{T}_1| = |\mathbf{W}_1| - |\mathbf{W}_2| = 50g - 15g = 35g \simeq 343.$$

◀ Interpret solution ▶

So the normal reaction of the floor on the sack has magnitude about 343 N ($35g$) and is directed upwards. In the absence of the stone and the pulley, the magnitude of the normal reaction would have been equal to the magnitude of the weight of the sack ($50g$). The effect of the stone, transmitted via the pulley, is as if the magnitude of the sack's weight were reduced by the magnitude of the weight of the stone ($15g$).

Note that the sack remains in contact with the floor if $|\mathbf{N}| > 0$, and is on the point of losing contact if $|\mathbf{N}|$ approaches zero. The model breaks down if the sack loses contact with the floor with $|\mathbf{N}| = 0$.

◀ Draw force diagram(s) ▶

(b) To answer this question, we need to model the forces on the pulley. We can consider the pulley as a particle of no mass. Let $\mathbf{T}_3$ and $\mathbf{T}_4$ represent the tension forces in the left- and right-hand ropes, respectively. Modelling the short piece of metal that attaches the pulley to the ceiling as a model string, with tension force $\mathbf{T}_5$, we have the force diagram shown in Figure 54.

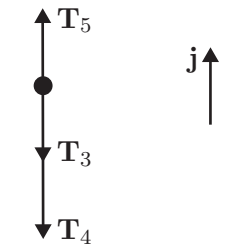


Figure 54 Force diagram for the pulley

◀ Apply law(s) ▶

The equilibrium condition for particles gives

$$\mathbf{T}_3 + \mathbf{T}_4 + \mathbf{T}_5 = \mathbf{0}. \tag{26}$$

Since the tension in a string around a model pulley remains constant, we have

$$|\mathbf{T}_1| = |\mathbf{T}_2| = |\mathbf{T}_3| = |\mathbf{T}_4|.$$

◀ Solve equation(s) ▶

Resolving equation (26) in the **j**-direction gives

$$-|\mathbf{T}_3| - |\mathbf{T}_4| + |\mathbf{T}_5| = 0.$$

Using the result from part (a) that $|\mathbf{T}_1| = |\mathbf{T}_2| = |\mathbf{W}_2|$, we have

$$|\mathbf{T}_5| = |\mathbf{T}_3| + |\mathbf{T}_4| = 2|\mathbf{W}_2| = 30g \simeq 294.$$

◀ Interpret solution ▶

So the model predicts that the force exerted by the short piece of metal (shown in Figure 52) on the pulley is about 294 N ($30g$) upwards. Hence, by Newton's third law, there is an equal and opposite force exerted by this short piece of metal on the ceiling – that is, the force exerted by the pulley on the ceiling is about 294 N ($30g$) downwards. This force (which is twice the weight of the stone) balances the weights of the stone ($15g$) and the sack ($50g$), less the normal reaction ($35g$) of the floor on the sack.

Exercise 35

Suppose that the pulley in Example 10 is no longer immediately above the sack, so that the rope attached to the sack makes an angle of $\pi/4$ to the vertical, as shown in Figure 55.

(a) Set up the problem using Procedure 2.

(b) Find the magnitude of the normal reaction of the floor on the sack.

(c) Find the magnitude of the friction force on the sack.

(d) Find the smallest value of the coefficient of static friction that would allow the system to remain in equilibrium.

(e) Find the weight of stone such that the sack just remains in contact with the floor.

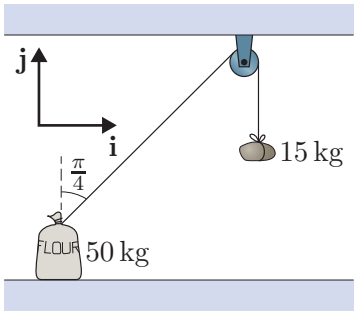


Figure 55

Slipping

You have already investigated slipping in the case of a one-particle system. In this subsection we examine the phenomenon in systems of more than one particle.

Example 11

Consider a scarf draped over the edge of a table. Model the scarf as two particles, one of mass m_1 hanging over the edge, and the other of mass m_2 resting on the table, with the masses joined by a model string passing over the edge of the table, which is modelled as a model pulley. Assume that the scarf's mass is uniformly distributed along its length, so that the masses of the two particles are proportional to the corresponding lengths of scarf.

If the coefficient of static friction between the scarf and the table surface is μ, what proportion of the scarf's length can hang over the edge of the table before the scarf slips off the table?

Solution

We can answer this question if we can find the ratio of m_1 (the mass of scarf hanging over the edge) to $m_1 + m_2$ (the total mass of scarf) when the scarf is on the verge of slipping.

The situation is illustrated in Figure 56, which also shows a suitable choice of axes.

◀ Draw picture ▶

◀ Choose axes ▶

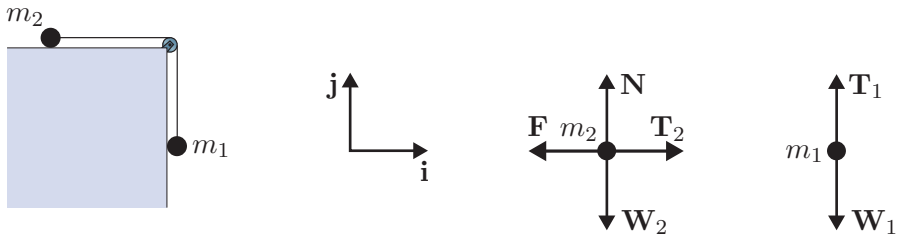


Figure 56 Modelling a scarf draped over the edge of a table, with force diagrams

There are two forces acting on the hanging particle: its weight $\mathbf{W}_1$ and the tension force $\mathbf{T}_1$. There are four forces acting on the particle on the table: its weight $\mathbf{W}_2$, the tension force $\mathbf{T}_2$, the normal reaction force $\mathbf{N}$, and the friction force $\mathbf{F}$. The force diagrams are shown in Figure 56.

◀ Draw force diagram(s) ▶

While the scarf does not slip, we can apply the equilibrium condition for particles to each particle in turn. For the hanging particle, we have

◀ Apply law(s) ▶

$$\mathbf{T}_1 + \mathbf{W}_1 = \mathbf{0}. \tag{27}$$

For the particle on the table, we have

$$\mathbf{F} + \mathbf{N} + \mathbf{T}_2 + \mathbf{W}_2 = \mathbf{0}. \tag{28}$$

The assumption of a model pulley gives

$$|\mathbf{T}_1| = |\mathbf{T}_2|. \tag{29}$$

When the particle is on the verge of slipping, we have

$$|\mathbf{F}| = \mu|\mathbf{N}|. \tag{30}$$

From Figure 56, the component forms of the force vectors can immediately be written down:

◀ Solve equation(s) ▶

$$\mathbf{T}_1 = |\mathbf{T}_1|\,\mathbf{j}, \quad \mathbf{W}_1 = |\mathbf{W}_1|\,(-\mathbf{j}),$$
$$\mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}), \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{T}_2 = |\mathbf{T}_2|\,\mathbf{i}, \quad \mathbf{W}_2 = |\mathbf{W}_2|\,(-\mathbf{j}).$$

Resolving equation (27) in the $\mathbf{j}$-direction gives

$$|\mathbf{T}_1| - |\mathbf{W}_1| = 0,$$

so

$$|\mathbf{T}_1| = |\mathbf{W}_1| = m_1 g.$$

Resolving equation (28) in the **i**-direction gives

$$-|\mathbf{F}| + 0 + |\mathbf{T}_2| + 0 = 0,$$

so, using equation (29),

$$|\mathbf{F}| = |\mathbf{T}_2| = |\mathbf{T}_1| = m_1 g.$$

Resolving equation (28) in the **j**-direction gives

$$0 + |\mathbf{N}| + 0 - |\mathbf{W}_2| = 0,$$

so

$$|\mathbf{N}| = |\mathbf{W}_2| = m_2 g.$$

Using equation (30), we have

$$m_1 g = \mu m_2 g,$$

so

$$m_1 = \mu m_2.$$

◀ Interpret solution ▶

Therefore the model predicts that when the scarf is on the verge of slipping, the proportion of its length that hangs over the edge is

$$\frac{m_1}{m_1 + m_2} = \frac{\mu m_2}{\mu m_2 + m_2} = \frac{\mu}{\mu + 1}.$$

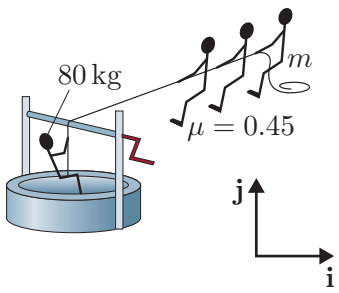


Figure 57 Lowering a man into a well

Exercise 36

A man of mass 80 kg is about to be lowered into a well from a rope that passes over a horizontal rotating axle above the well. The other end of the rope is held by several men each of mass 80 kg, as shown in Figure 57, with the rope between the men and the axle horizontal. Assume that n men can be represented by a single particle of mass $80n$ kg, the rope by a model string and the axle by a model pulley.

If the coefficient of static friction between the men's boots and the ground is 0.45, how many men are required to hold the man at the end of the rope before he is lowered into the well?

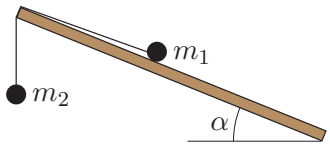


Figure 58 Two connected particles on an inclined plane

Exercise 37

An object of mass m_1, resting on a board inclined at an angle α to the horizontal, is attached to an object of mass m_2 by a string hanging over the edge of the board, as shown in Figure 58.

Assuming that the objects can be modelled as particles, the string as a model string and the edge of the board as a model pulley, find the condition on the coefficient of static friction μ between the first object and the board for this system to remain in equilibrium.

(*Hint*: There are two ways in which the equilibrium can be disturbed.)

4 Torques

This section looks at solid bodies, and in particular at a phenomenon that does not apply to particles: the turning effect of forces. We begin in Subsection 4.1 by introducing ways of modelling objects when their size is important, as it is when the turning effects of forces are considered. Subsection 4.2 goes on to explain what is meant by the turning effect of a force, and to provide a mathematical description of such an effect.

4.1 Extended and rigid bodies

Consider, for example, a tall thin box of cereal on a breakfast table. If you push it near its base, it slides across the table. But if you push it near its top, it tips over. The *position* at which the pushing force acts is important here, so the particle model – which allows forces to act at only one point – is inadequate. In this situation, we model the cereal box as an **extended body**, which is defined to be a material object that has one or more of length, breadth and depth, but whose internal structure may be neglected. So like a particle, it has mass, but unlike a particle, it has size of some sort and occupies more than a single point in space. For a particle, all the forces act at a point. For a rigid body, the forces can act at different points on the body.

The shape of an extended body may be one-, two- or three-dimensional (i.e. it may have a length, an area or a volume), depending on the situation.

Extended bodies are complicated objects because they can flex or vibrate. To model the slipping or tipping behaviour of the cereal packet, this generality is not needed. So we restrict our attention to *rigid bodies*. A **rigid body** is defined to be an extended body that does not change its shape (so it does not flex or vibrate).

The formal definition of a rigid body states that the position vector of a point on the body relative to any other point on the body is constant.

For a particle, all forces are applied at the point represented by the particle; we say that the forces *act* at this point. For extended bodies, we need to be more careful to specify where a force acts. For example, the weight of a body always acts through the **centre of mass** of the body. In this unit we consider only symmetric bodies made of uniform material. For such a body, as you would intuitively expect, the centre of mass is at its centre of symmetry (or geometric centre). For example, consider a coin, which can be modelled as a disc. Using symmetry, we can state that the centre of mass of the coin is along the axis of the disc (which runs between the centres of the flat circular faces of the disc), halfway between the flat circular faces.

A definition of the centre of mass and ways of finding it for a general rigid body are given in Units 19 and 21.

4.2 Turning effect of a force

Suppose that you try to balance a ruler on a horizontal extended finger. When the centre of the ruler is over your finger, the ruler should balance. If the centre of the ruler is not over your finger, then the weight of the ruler will cause it to turn about your finger. Although the finger can provide a normal reaction force that is equal in magnitude to the weight of

the ruler, if the two forces are not in line, then turning occurs. The weight provides a turning effect if the vertical line of its action (through the centre of mass) does not pass through your finger.

How do we measure the turning effect of a force in terms of what we already know, namely the magnitude and direction of the force, and the point on the object at which the force acts? In order to begin to answer this question, try a simple experiment. Balance a 30-centimetre (12-inch) ruler on a pencil rubber or a hexagonal pencil, or some other object that is not too wide and so will act as a pivot. Place two identical small coins on either side of the pivot so that each is 10 cm from the pivot (see Figure 59).

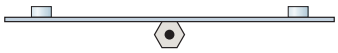


Figure 59 A ruler balanced on a pencil, with coins at each end

Then experiment with moving one of the coins in steps of 2 cm from its initial position, and see how the other coin has to be moved in order to re-establish balance. The conclusion from this experiment is not, perhaps, a surprising one: coins of equal mass have to be placed at equal distances on either side of the pivot for the ruler to remain balanced.

Next place the two coins together at a point on the ruler, say at 6 cm from the pivot. Where does a third identical coin have to be placed to achieve balance?

You should find that two identical coins placed together at 6 cm from the pivot are balanced by another identical coin placed on the other side of the ruler at 12 cm from the pivot. If you continue to experiment with varying numbers of coins placed at various pairs of positions along the ruler, you will find in each case that if the masses of the two sets of coins are unequal, then in order to achieve balance, the greater mass has to be placed nearer to the pivot than the smaller one. The turning effect due to the weight of the coins acting at a point depends not only on the mass of the coins, but also on the distance of the point of action from the pivot.

A long symmetrical object, such as the ruler in this experiment, can often be modelled as a rigid body with length, but no breadth or depth. Such a rigid body is known as a **model rod** and is often drawn as a straight line. The pivot on which the object rests is often modelled as a **model pivot**, which has a single point of contact with the rod and is often drawn as a triangle. Using these notions, the above experiments should allow you to believe the following result.

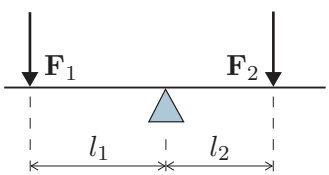


Figure 60 A balanced rod

Balanced rod

A horizontal rod, pivoted at its centre, will remain horizontal under the action of two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ acting vertically downwards at distances l_1 and l_2, respectively, on either side of the pivot (see Figure 60), provided that

$$|\mathbf{F}_1|\, l_1 = |\mathbf{F}_2|\, l_2. \tag{31}$$

In other words, the horizontal rod will remain in equilibrium provided that the distances of the forces from the pivot are in inverse ratio to the magnitudes of the forces.

Exercise 38

Jack and Jill are sitting on opposite ends of a seesaw. Jill is 1.2 m from the pivot, and Jack is 1 m from the pivot. Jack's mass is 60 kg. If the seesaw is at rest and horizontal, what is Jill's mass?

In the situation described in Exercise 38 and in the ruler example, there was an obvious way to measure each distance, namely along the seesaw or ruler. We need to generalise this to situations where there is not such an obvious way to measure distance. In these two examples the distances along the ruler and seesaw happen to be distances measured perpendicular to the direction of the force and from its point of action, that is, perpendicular to the *line of action* of the force.

The **line of action of a force** is a straight line in the direction of the force and through the point of action of the force.

So the turning effect of a force about a fixed point needs to encompass a measure of the force itself and the perpendicular distance of its line of action from the fixed point. It also needs to increase in magnitude if either the force or the distance increases in magnitude, and vice versa.

Consider a force $\mathbf{F}$ with line of action AB, as shown in Figure 61, and some fixed point O. Let R be any point on AB with position vector $\mathbf{r}$ with respect to O. Then the cross product is

$$\mathbf{r} \times \mathbf{F} = (|\mathbf{r}|\,|\mathbf{F}| \sin\theta)\,\hat{\mathbf{n}},$$

where $\hat{\mathbf{n}}$ is a unit vector perpendicular to both $\mathbf{r}$ and $\mathbf{F}$, and with direction out of the page (as given by the right-hand grip rule). This cross product satisfies all the above requirements for the turning effect of a force. It includes a measure $|\mathbf{F}|$ of the force and the perpendicular distance $|\mathbf{r}| \sin\theta$ of its line of action AB from the fixed point O. The direction of the cross product, represented by $\hat{\mathbf{n}}$, corresponds to the direction of the turning effect. In this example, the turning effect of $\mathbf{F}$ about O is anticlockwise, which corresponds to the anticlockwise motion of the fingers in the right-hand grip rule. We refer to the cross product $\mathbf{r} \times \mathbf{F}$ as the *torque* of the force $\mathbf{F}$ relative to the origin O, and we use it as our measure of the turning effect of the force.

Here $\mathbf{F}$ is used to denote a general force rather than a friction force.

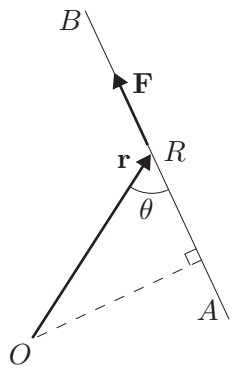


Figure 61 Calculating the turning effect of force $\mathbf{F}$ relative to O

The **torque** $\boldsymbol{\Gamma}$ of a force $\mathbf{F}$ about a fixed point O is the cross product

$$\boldsymbol{\Gamma} = \mathbf{r} \times \mathbf{F},$$

where $\mathbf{r}$ is the position vector, relative to O, of a point on the line of action of the force.

In SI units, the units of torque are newton metres, written as N m (or $\mathrm{kg\,m^2\,s^{-2}}$).

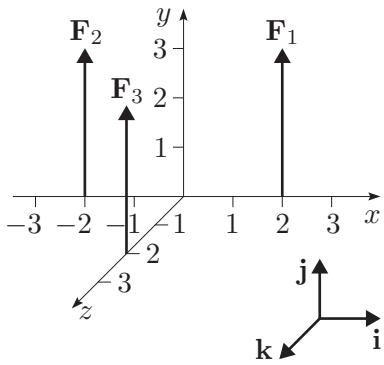


Figure 62

Note that a torque is a vector quantity, but it does not have the same units as a force, which is measured in newtons.

Exercise 39

Find the torque of each of the forces in Figure 62 relative to the origin, where each force is of magnitude 3 N. Forces $\mathbf{F}_1$ and $\mathbf{F}_2$ each have their point of action on the x-axis, and $\mathbf{F}_3$ has its point of action on the z-axis.

Exercise 40

(a) Show that if O is any point on the line of action of a force $\mathbf{F}$, and $\mathbf{r}$ is the position vector, relative to O, of any other point on the line of action, then $\mathbf{r} \times \mathbf{F} = \mathbf{0}$. Deduce that the torque of a force about a point on its line of action is zero.

(b) Suppose now that O is *not* on the line of action of $\mathbf{F}$, and let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the position vectors, relative to O, of two points on the line of action. Show that $\mathbf{r}_1 - \mathbf{r}_2$ is parallel to $\mathbf{F}$, and hence that $\mathbf{r}_1 \times \mathbf{F} = \mathbf{r}_2 \times \mathbf{F}$. Deduce that the torque of a force about a fixed point O is independent of the choice of the position of the point on the line of action.

Let us now convince ourselves that the definition of torque makes sense in terms of the examples of turning forces that we have seen so far. To do this, we need an equilibrium condition for rigid bodies that extends the equilibrium condition for particles. You will not be surprised that it requires not only that all the forces on a rigid body sum to zero, but also that all the torques sum to zero.

Equilibrium condition for a rigid body

A rigid body subjected to forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ is in **equilibrium** if the forces sum to the zero vector and the torques $\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \ldots, \boldsymbol{\Gamma}_n$ corresponding to the forces, relative to the same fixed point O, also sum to the zero vector, that is,

$$\sum_{i=1}^{n} \mathbf{F}_i = \mathbf{0} \quad \text{and} \quad \sum_{i=1}^{n} \boldsymbol{\Gamma}_i = \sum_{i=1}^{n} \mathbf{r}_i \times \mathbf{F}_i = \mathbf{0}, \tag{32}$$

where $\mathbf{r}_i$ is the position vector, relative to O, of a point on the line of action of $\mathbf{F}_i$.

5 Applying the principles

This section contains some examples and exercises that use the principles developed so far to solve more complicated statics problems. These examples are more complicated than those in the previous section because the forces and positions are not conveniently aligned. This is not a fundamental difficulty; it merely makes the calculation of the vector products more complicated.

Example 12

A ladder of mass M and length l stands on rough horizontal ground, and rests against a smooth vertical wall (see Figure 63). The ladder can be modelled as a model rod. Find the minimum angle θ between the ladder and the ground for which the ladder can remain static, if the coefficient of static friction μ between the ladder and the ground is 0.5.

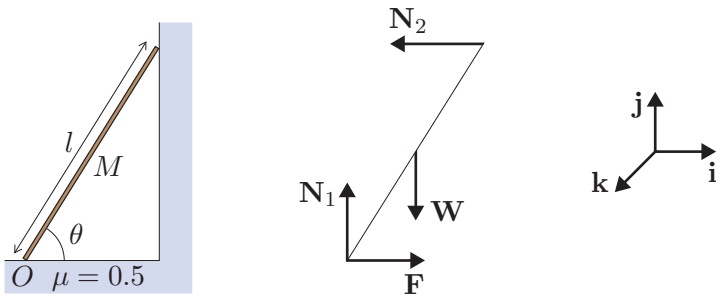


Figure 63 A ladder resting against a wall

Solution

The forces acting on the rod are shown in the force diagram in Figure 63, where $\mathbf{W}$ is the weight of the ladder, $\mathbf{N}_1$ is the normal reaction from the ground, $\mathbf{N}_2$ is the normal reaction from the wall, and $\mathbf{F}$ is the friction force at the bottom of the ladder. (The wall is smooth, so there is no friction force at the top of the ladder.)

◀ Draw force diagram(s) ▶

As two of the three unknown forces act at the bottom of the ladder, this is a convenient point for the origin O. The axes are chosen as shown in Figure 63.

◀ Choose axes ▶

The equilibrium condition for rigid bodies gives

◀ Apply law(s) ▶

$$\mathbf{N}_1 + \mathbf{N}_2 + \mathbf{F} + \mathbf{W} = \mathbf{0}, \tag{33}$$

$$\boldsymbol{\Gamma}_{\mathrm{N}_1} + \boldsymbol{\Gamma}_{\mathrm{N}_2} + \boldsymbol{\Gamma}_{\mathrm{F}} + \boldsymbol{\Gamma}_{\mathrm{W}} = \mathbf{0}, \tag{34}$$

where $\boldsymbol{\Gamma}_{\mathrm{N}_1}$, $\boldsymbol{\Gamma}_{\mathrm{N}_2}$, $\boldsymbol{\Gamma}_{\mathrm{F}}$ and $\boldsymbol{\Gamma}_{\mathrm{W}}$ are the torques with respect to O of $\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{F}$ and $\mathbf{W}$, respectively.

If the ladder is not going to slip, we must have

$$|\mathbf{F}| \leq \mu |\mathbf{N}_1|. \tag{35}$$

◀ Solve equation(s) ▶

The position vectors $\mathbf{r}_{N_1}$ and $\mathbf{r}_F$ are both zero, hence the corresponding torques $\boldsymbol{\Gamma}_{N_1}$ and $\boldsymbol{\Gamma}_F$ (relative to O) are also zero. To calculate the non-zero torques, we need the position vectors of the points of application of the forces. These are given by

$$\mathbf{r}_{N_2} = l\cos\theta\,\mathbf{i} + l\sin\theta\,\mathbf{j},$$
$$\mathbf{r}_W = \tfrac{1}{2}l\cos\theta\,\mathbf{i} + \tfrac{1}{2}l\sin\theta\,\mathbf{j}.$$

All of the forces in this example are aligned with the coordinate axes, so the components can be written down by inspection:

$$\mathbf{N}_1 = |\mathbf{N}_1|\,\mathbf{j}, \quad \mathbf{N}_2 = |\mathbf{N}_2|\,(-\mathbf{i}), \quad \mathbf{F} = |\mathbf{F}|\,\mathbf{i}, \quad \mathbf{W} = Mg\,(-\mathbf{j}).$$

Now we can calculate the two non-zero torques:

$$\begin{aligned}\boldsymbol{\Gamma}_{N_2} &= (l\cos\theta\,\mathbf{i} + l\sin\theta\,\mathbf{j}) \times (-|\mathbf{N}_2|\,\mathbf{i}) \\ &= -|\mathbf{N}_2|\,l\sin\theta\,\mathbf{j}\times\mathbf{i} \\ &= |\mathbf{N}_2|\,l\sin\theta\,\mathbf{k},\end{aligned}$$

$$\begin{aligned}\boldsymbol{\Gamma}_W &= (\tfrac{1}{2}l\cos\theta\,\mathbf{i} + \tfrac{1}{2}l\sin\theta\,\mathbf{j}) \times (-Mg\,\mathbf{j}) \\ &= -\tfrac{1}{2}Mgl\cos\theta\,\mathbf{i}\times\mathbf{j} \\ &= -\tfrac{1}{2}Mgl\cos\theta\,\mathbf{k}.\end{aligned}$$

Substituting these torques into equation (34) gives

$$|\mathbf{N}_2|\,l\sin\theta\,\mathbf{k} - \tfrac{1}{2}Mgl\cos\theta\,\mathbf{k} = \mathbf{0},$$

thus

$$|\mathbf{N}_2| = \tfrac{1}{2}Mg\cot\theta.$$

Resolving equation (33) in the **i**- and **j**-directions in turn gives

$$-|\mathbf{N}_2| + |\mathbf{F}| = 0, \quad |\mathbf{N}_1| - Mg = 0.$$

Therefore

$$|\mathbf{F}| = |\mathbf{N}_2| = \tfrac{1}{2}Mg\cot\theta, \quad |\mathbf{N}_1| = Mg.$$

Substituting these into inequality (35) gives

$$\tfrac{1}{2}Mg\cot\theta \le \mu Mg,$$

Note that since M and g are positive, we can safely divide through by them without reversing the inequality.

which, on rearrangement and using $\mu = 0.5$, gives

$$\cot\theta \le 1.$$

◀ Interpret results ▶

Therefore the model predicts that the minimum angle that the ladder can make with the ground before slipping is $\pi/4$ radians (45°).

The second example involves a *hinge*, where a body is constrained to rotate about a fixed point. For such problems we need a new type of force, a **reaction force R**, which models the force acting on a hinge.

Example 13

During the erection of a marquee, a heavy pole OA of mass m and length l must be held in place by a rope AB, as shown in Figure 64. The angle between the pole and the ground is $\pi/4$, and the angle between the rope and the ground is $\pi/6$. Model the pole as a model rod, and the rope as a model string. Assume that the pole is freely hinged at O, that is, the end of the model rod is fixed at O and is free to pivot about O.

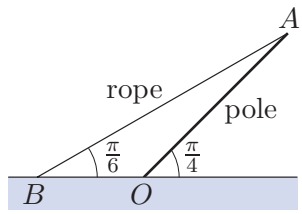


Figure 64 A rope holding up a pole

If the pole is in equilibrium, find the magnitude of the tension in the rope. Is the magnitude of the tension in the rope larger or smaller than it would be if the pole were hanging freely on the end of the rope?

Solution

◀ Draw force diagram(s) ▶

The best choice of axes is not obvious for this problem. So in this case we proceed by drawing the force diagram first, as shown in Figure 65.

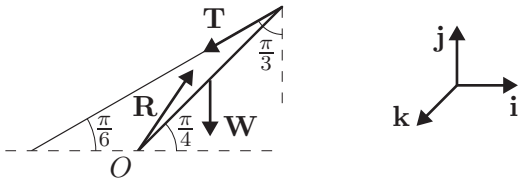


Figure 65 The force diagram for the rope and pole

◀ Choose axes ▶

From Figure 65 we see that there are three forces acting on the rod: its weight $\mathbf{W}$, the tension force in the string $\mathbf{T}$, and the reaction force $\mathbf{R}$ at the hinge. These forces act in three different directions, so we choose to orient the axes as shown in the figure. In this problem, the choice of origin O is vital, since we are going to take torques about the origin. Choosing the origin at the base of the pole eliminates the reaction force from the torque equation, which simplifies the calculations.

◀ Apply law(s) ▶

The equilibrium condition for rigid bodies gives

$$\mathbf{R} + \mathbf{W} + \mathbf{T} = \mathbf{0}, \tag{36}$$

$$\boldsymbol{\Gamma}_{\mathrm{R}} + \boldsymbol{\Gamma}_{\mathrm{W}} + \boldsymbol{\Gamma}_{\mathrm{T}} = \mathbf{0}, \tag{37}$$

where $\boldsymbol{\Gamma}_{\mathrm{R}}$, $\boldsymbol{\Gamma}_{\mathrm{W}}$ and $\boldsymbol{\Gamma}_{\mathrm{T}}$ are the torques with respect to O of $\mathbf{R}$, $\mathbf{W}$ and $\mathbf{T}$, respectively.

◀ Solve equation(s) ▶

In this case, equation (36) is not very useful since it contains two forces of unknown magnitude, namely $\mathbf{R}$ and $\mathbf{T}$. However, by taking torques about O, one of the torques appearing in equation (37) is zero, namely $\boldsymbol{\Gamma}_{\mathrm{R}}$. Now we proceed by calculating the other two torques.

The position vectors of the points of application of the forces $\mathbf{W}$ and $\mathbf{T}$ (at the centre and end of the pole, respectively) are

$$\mathbf{r}_{\mathrm{W}} = \tfrac{1}{2}l\cos\tfrac{\pi}{4}\,\mathbf{i} + \tfrac{1}{2}l\sin\tfrac{\pi}{4}\,\mathbf{j} = \frac{l}{2\sqrt{2}}(\mathbf{i}+\mathbf{j}),$$

$$\mathbf{r}_{\mathrm{T}} = l\cos\tfrac{\pi}{4}\,\mathbf{i} + l\sin\tfrac{\pi}{4}\,\mathbf{j} = \frac{l}{\sqrt{2}}(\mathbf{i}+\mathbf{j}).$$

There are many ways of computing the components of vectors. It is up to you to choose the way that you feel most comfortable with.

The weight of the pole can be written in component form by inspection of the diagram, as $\mathbf{W} = mg\,(-\mathbf{j})$. The angle between the direction of the tension force and the vertical is shown in Figure 65 to be $\pi/3$. We have

$$\mathbf{T} = |\mathbf{T}|\sin\tfrac{\pi}{3}\,(-\mathbf{i}) + |\mathbf{T}|\cos\tfrac{\pi}{3}\,(-\mathbf{j}) = -|\mathbf{T}|\left(\tfrac{\sqrt{3}}{2}\,\mathbf{i} + \tfrac{1}{2}\,\mathbf{j}\right).$$

Now that the forces and position vectors are written in component form, we can proceed to calculate the torques:

$$\begin{aligned}\boldsymbol{\Gamma}_{\mathrm{W}} = \mathbf{r}_{\mathrm{W}} \times \mathbf{W} &= \frac{l}{2\sqrt{2}}(\mathbf{i}+\mathbf{j}) \times (-mg\,\mathbf{j}) \\ &= -\frac{lmg}{2\sqrt{2}}\,\mathbf{k},\end{aligned}$$

$$\begin{aligned}\boldsymbol{\Gamma}_{\mathrm{T}} = \mathbf{r}_{\mathrm{T}} \times \mathbf{T} &= \frac{l}{\sqrt{2}}(\mathbf{i}+\mathbf{j}) \times \left(-|\mathbf{T}|\left(\tfrac{\sqrt{3}}{2}\,\mathbf{i} + \tfrac{1}{2}\,\mathbf{j}\right)\right) \\ &= -\frac{l\,|\mathbf{T}|}{\sqrt{2}}\left(\tfrac{1}{2}\,\mathbf{i}\times\mathbf{j} + \tfrac{\sqrt{3}}{2}\,\mathbf{j}\times\mathbf{i}\right) \\ &= \frac{l\,|\mathbf{T}|}{2\sqrt{2}}(\sqrt{3}-1)\mathbf{k}.\end{aligned}$$

Substituting these torques into equation (37) gives

$$-\frac{lmg}{2\sqrt{2}}\,\mathbf{k} + \frac{l\,|\mathbf{T}|}{2\sqrt{2}}(\sqrt{3}-1)\mathbf{k} = \mathbf{0}.$$

Rearranging gives

$$|\mathbf{T}| = \frac{mg}{\sqrt{3}-1}.$$

◀ Interpret results ▶

So the magnitude of the tension in the rope is $mg/(\sqrt{3}-1) \simeq 1.4mg$, which is greater than the magnitude of the weight of the pole, mg. So, rather counter-intuitively, a stronger rope is needed to erect a pole in this way than is needed to lift the pole.

Exercise 41

A model rod OA of length l and mass m is fixed to a wall by a free hinge, as shown in Figure 66. The rod is free to turn in a vertical plane about the hinge, which is assumed to be smooth (i.e. there is no friction force associated with it). The rod is supported in a horizontal position by a string AB inclined at an angle θ to the horizontal.

Find the reaction force at the hinge and the tension force acting on the rod due to the string. Comment on the magnitudes and directions of these two forces.

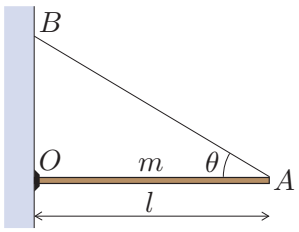


Figure 66 A rod fixed to a wall by a free hinge and supported by a string

Exercise 42

A light ladder of length 3.9 m stands on a horizontal floor and rests against a smooth vertical wall, as shown in Figure 67.

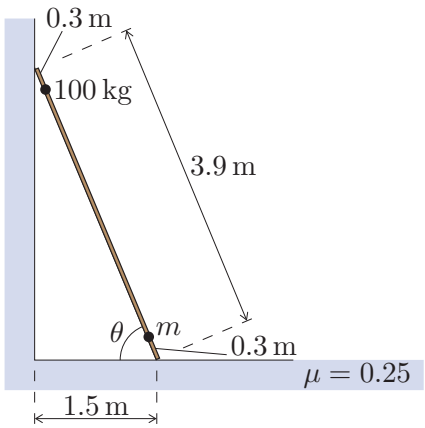


Figure 67 A ladder with people standing on the top and bottom rungs

The base of the ladder is 1.5 m from the base of the wall. The coefficient of static friction between the ladder and the floor is 0.25. The end rungs are each 0.3 m from an end of the ladder. The ladder may be modelled as a rod, and its mass may be neglected (as it is a light ladder, so its mass is negligible compared with the masses of any people standing on it).

What is the minimum mass of a person standing on the bottom rung that prevents the ladder from slipping when a person of mass 100 kg stands on the top rung?

Exercise 43

At a building site, a plank OB of length $2l$ and mass m is resting against a large smooth pipe of radius r, as shown in Figure 68. The pipe is fixed to the ground, or else it would slide or roll to the left. The angle between the plank and the horizontal is $\pi/3$, so by symmetry the angle between the horizontal and the line between O and the centre of the pipe is $\pi/6$, as shown. In the figure, the distance OA is greater than the distance AB, so the centre of mass of the plank is between O and A.

What is the least coefficient of static friction between the plank and the ground that will ensure equilibrium?

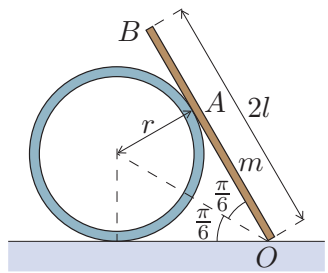


Figure 68 A plank resting against a large smooth pipe

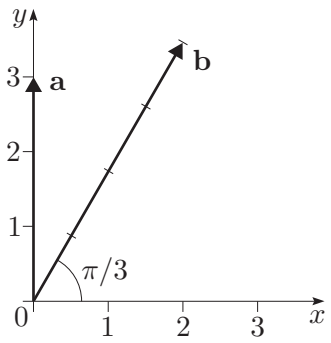
Learning outcomes

After studying this unit, you should be able to:

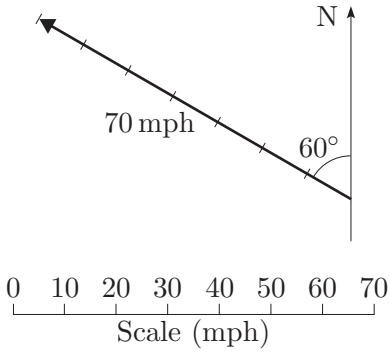
- understand the meaning of the terms scalar, vector, displacement vector, unit vector and position vector, and know what it means to say that two vectors are equal
- use vector notation and represent vectors as arrows on diagrams
- scale a vector by a number, and add two vectors geometrically using the triangle rule (or the parallelogram rule)
- resolve a vector into its Cartesian components, and scale and add vectors given in Cartesian component form
- calculate the dot product and cross product of two vectors
- determine whether or not two given vectors are perpendicular or parallel to one another
- determine the magnitude of a vector and the angle between the directions of two vectors
- write down the vector equation of a given straight line
- resolve a vector in a given direction
- manipulate vector expressions and equations involving the scaling, addition, dot product and cross product of vectors
- use the cross product to determine the area of a parallelogram or triangle
- appreciate the concept of a force
- understand and model forces of weight, normal reaction, tension and friction
- recognise and model the forces that act on an object in equilibrium
- model objects as particles or as rigid bodies
- use model strings, model rods, model pulleys and model pivots in modelling systems involving forces
- draw force diagrams, and choose appropriate axes and an origin
- use the equilibrium conditions for particles and for rigid bodies
- understand and use torques
- model and solve a variety of problems involving systems in equilibrium and systems on the verge of leaving equilibrium.

Solutions to exercises

Solution to Exercise 1



Solution to Exercise 2



Solution to Exercise 3

(a) $-\mathbf{v}$ represents a wind of 35 mph from the south-west.

(b) The vector $-1.5\mathbf{v}$ has magnitude $1.5|\mathbf{v}|$ and direction opposite to $\mathbf{v}$.

The vector $-k\mathbf{v}$ (k positive) has magnitude $k|\mathbf{v}|$ and direction opposite to $\mathbf{v}$.

(c) $\dfrac{1}{|\mathbf{v}|}\mathbf{v}$ is a scaling of $\mathbf{v}$ by the positive scalar $m = \dfrac{1}{|\mathbf{v}|}$. The direction of $\dfrac{1}{|\mathbf{v}|}\mathbf{v}$ is thus the same as that of $\mathbf{v}$, and its magnitude is

$$m|\mathbf{v}| = \frac{1}{|\mathbf{v}|}|\mathbf{v}| = 1.$$

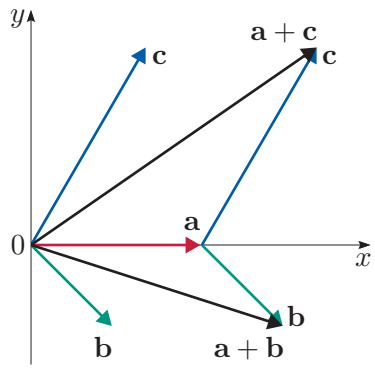
Solution to Exercise 4

(a) $-35\,\mathbf{j}$ (where $|\mathbf{j}|$ represents 1 km per hour).

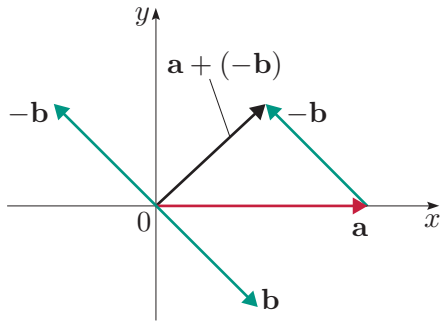
(b) $-112\,\mathbf{i}$ (where $|\mathbf{i}|$ represents 1 mile).

Solution to Exercise 5

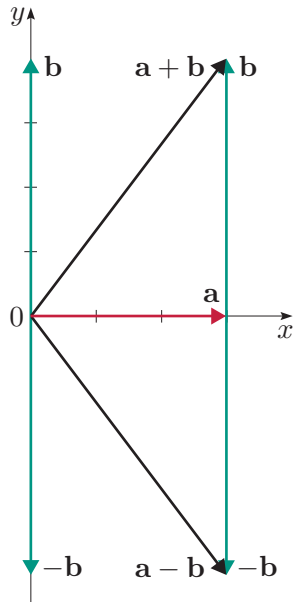
(a) The figure below shows $\mathbf{a}+\mathbf{b}$ and $\mathbf{a}+\mathbf{c}$.



(b) The figure below shows $-\mathbf{b}$ and the sum of $\mathbf{a}$ and $-\mathbf{b}$.



Solution to Exercise 6



Solution to Exercise 7

$$4(\mathbf{a}-\mathbf{c})+3(\mathbf{c}-\mathbf{b})+2(2\mathbf{a}-\mathbf{b}-3\mathbf{c})$$
$$=4\mathbf{a}-4\mathbf{c}+3\mathbf{c}-3\mathbf{b}+4\mathbf{a}-2\mathbf{b}-6\mathbf{c}$$
$$=8\mathbf{a}-5\mathbf{b}-7\mathbf{c}.$$

Solution to Exercise 8

Systems (b), (c) and (d) are right-handed.

Solution to Exercise 9

(a) $\mathbf{d}=2(\mathbf{i}+\mathbf{j}+\mathbf{k})-3(2\mathbf{i}-3\mathbf{j}-\mathbf{k})=-4\mathbf{i}+11\mathbf{j}+5\mathbf{k}$,

$\mathbf{e}=(\mathbf{i}+\mathbf{j}+\mathbf{k})-2(2\mathbf{i}-3\mathbf{j}-\mathbf{k})+4(3\mathbf{i}+\mathbf{k})=9\mathbf{i}+7\mathbf{j}+7\mathbf{k}$.

(b) $|\mathbf{d}|=\sqrt{(-4)^2+11^2+5^2}=\sqrt{162}=9\sqrt{2}$,

$|\mathbf{e}|=\sqrt{9^2+7^2+7^2}=\sqrt{179}$.

(c) $|\mathbf{a}|=\sqrt{1^2+1^2+1^2}=\sqrt{3}$.

A unit vector in the direction of $\mathbf{a}$ is

$$\frac{1}{|\mathbf{a}|}\mathbf{a}=\frac{1}{\sqrt{3}}(\mathbf{i}+\mathbf{j}+\mathbf{k}).$$

(d) If $\mathbf{a}+\mathbf{x}=\mathbf{b}$, then

$$\mathbf{x}=\mathbf{b}-\mathbf{a}=(2\mathbf{i}-3\mathbf{j}-\mathbf{k})-(\mathbf{i}+\mathbf{j}+\mathbf{k})=\mathbf{i}-4\mathbf{j}-2\mathbf{k}.$$

Thus the components of $\mathbf{x}$ are 1, -4 and -2.

Solution to Exercise 10

$\mathbf{0}=(0\quad 0\quad 0)^T$, $\mathbf{i}=(1\quad 0\quad 0)^T$, $\mathbf{j}=(0\quad 1\quad 0)^T$, $\mathbf{k}=(0\quad 0\quad 1)^T$.

Solution to Exercise 11

(a) Relative to the origin of the Cartesian coordinate system, the two points have position vectors $\mathbf{i}+\mathbf{j}+2\mathbf{k}$ and $2\mathbf{i}+3\mathbf{j}+\mathbf{k}$. Thus the vector equation of the line is

$$\mathbf{r}=(1-s)(\mathbf{i}+\mathbf{j}+2\mathbf{k})+s(2\mathbf{i}+3\mathbf{j}+\mathbf{k})$$
$$=(1+s)\mathbf{i}+(1+2s)\mathbf{j}+(2-s)\mathbf{k},$$

where $-\infty<s<\infty$.

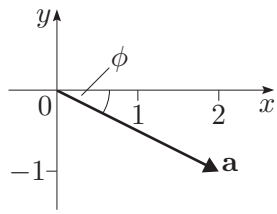
(b) The line cuts the (x,z)-plane when $y=0$, that is, $1+2s=0$, giving $s=-\frac{1}{2}$. Hence $\mathbf{r}=\frac{1}{2}\mathbf{i}+\frac{5}{2}\mathbf{k}$.

Solution to Exercise 12

(a) $|\mathbf{a}|=\sqrt{2^2+(-1)^2}=\sqrt{5}$,

$|\mathbf{b}|=\sqrt{1^2+3^2+5^2}=\sqrt{35}$.

The vector $\mathbf{a}$ lies in the (x,y)-plane, and the angle ϕ that it makes with the x-axis (see the figure in the margin) is given by $\cos\phi=2/\sqrt{5}$ and $\sin\phi=-1/\sqrt{5}$. Hence $\phi\simeq-0.4636$ (radians).



(b) $\mathbf{a}+\mathbf{b}=3\mathbf{i}+2\mathbf{j}+5\mathbf{k}$,

$2\mathbf{a}-\mathbf{b}=3\mathbf{i}-5\mathbf{j}-5\mathbf{k}$,

$\mathbf{c}+2\mathbf{b}-3\mathbf{a}=-3\mathbf{i}+10\mathbf{j}+8\mathbf{k}$.

(c) The vector $\overrightarrow{PQ}$ is equal to $2\mathbf{a}-\mathbf{b}$. The point Q is the end of the vector $\overrightarrow{OQ}$, which is given by

$$\overrightarrow{OQ}=\overrightarrow{OP}+\overrightarrow{PQ}=(2\mathbf{j}+3\mathbf{k})+(3\mathbf{i}-5\mathbf{j}-5\mathbf{k})=3\mathbf{i}-3\mathbf{j}-2\mathbf{k},$$

so Q is the point $(3,-3,-2)$.

Solution to Exercise 13

(a) Writing

$$\frac{x+1}{3}=\frac{y-8}{-2}=1-z=s$$

gives

$$x=3s-1,\quad y=8-2s,\quad z=1-s.$$

This gives the vector equation of the line as

$$\begin{aligned}\mathbf{r}&=(3s-1)\mathbf{i}+(8-2s)\mathbf{j}+(1-s)\mathbf{k}\\&=(-\mathbf{i}+8\mathbf{j}+\mathbf{k})+s(3\mathbf{i}-2\mathbf{j}-\mathbf{k}).\end{aligned}$$

Hence A has position vector $\mathbf{a}=-\mathbf{i}+8\mathbf{j}+\mathbf{k}$, B has position vector $\mathbf{b}=2\mathbf{i}+6\mathbf{j}$, and $\overrightarrow{AB}=3\mathbf{i}-2\mathbf{j}-\mathbf{k}$, with length $\sqrt{3^2+2^2+1^2}=\sqrt{14}$.

(b) (i) The midpoint between A and B has $s=\frac{1}{2}$, so M has position vector

$$\mathbf{m}=\tfrac{1}{2}\mathbf{i}+7\mathbf{j}+\tfrac{1}{2}\mathbf{k}.$$

(ii) The point N has $s=0.6$, so N has position vector

$$\mathbf{n}=0.8\mathbf{i}+6.8\mathbf{j}+0.4\mathbf{k}.$$

(c) Let $\mathbf{c}=-\mathbf{i}-2\mathbf{j}$ and $\mathbf{d}=2\mathbf{i}+\mathbf{j}-\mathbf{k}$. Then $\mathbf{d}-\mathbf{c}=3\mathbf{i}+3\mathbf{j}-\mathbf{k}$, so the vector equation of the line (using a different parameter t) is

$$\begin{aligned}\mathbf{r}=\mathbf{c}+t(\mathbf{d}-\mathbf{c})&=(-\mathbf{i}-2\mathbf{j})+t(3\mathbf{i}+3\mathbf{j}-\mathbf{k})\\&=(3t-1)\mathbf{i}+(3t-2)\mathbf{j}-t\mathbf{k}.\end{aligned}$$

(d) The lines intersect if there are values of s and t such that

$$(3t-1)\mathbf{i}+(3t-2)\mathbf{j}-t\mathbf{k}=(3s-1)\mathbf{i}+(8-2s)\mathbf{j}+(1-s)\mathbf{k}.$$

This gives rise to the three equations

$$3t-1=3s-1,\quad 3t-2=8-2s,\quad -t=1-s.$$

Since there are no values of s and t that satisfy all three equations, the lines do not intersect.

Solution to Exercise 14

$\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}|\,|\mathbf{b}|\cos\theta=2\times4\times\cos\frac{\pi}{3}=4$,

$\mathbf{b}\cdot\mathbf{c}=|\mathbf{b}|\,|\mathbf{c}|\cos\theta=4\times1\times\cos\frac{\pi}{6}=2\sqrt{3}$,

$\mathbf{a} \cdot \mathbf{c} = |\mathbf{a}|\,|\mathbf{c}| \cos\theta = 2 \times 1 \times \cos\left(\frac{\pi}{3} + \frac{\pi}{6}\right) = 2\cos\frac{\pi}{2} = 0,$

$\mathbf{b} \cdot \mathbf{b} = |\mathbf{b}|\,|\mathbf{b}| \cos\theta = 4 \times 4 \times \cos 0 = 16.$

Solution to Exercise 15

(a) $$\begin{aligned}(\mathbf{a}+\mathbf{b}) \cdot (\mathbf{a}-\mathbf{b}) &= \mathbf{a}\cdot\mathbf{a} - \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a} - \mathbf{b}\cdot\mathbf{b} \\ &= \mathbf{a}\cdot\mathbf{a} - \mathbf{b}\cdot\mathbf{b}.\end{aligned}$$

(b) $$\begin{aligned}|\mathbf{a}+\mathbf{b}|^2 &= (\mathbf{a}+\mathbf{b}) \cdot (\mathbf{a}+\mathbf{b}) \\ &= \mathbf{a}\cdot\mathbf{a} + \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} \\ &= \mathbf{a}\cdot\mathbf{a} + 2\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{b}.\end{aligned}$$

Solution to Exercise 16

(a) If $2\mathbf{a} + 3\mathbf{b}$ and $m\mathbf{a} + \mathbf{b}$ are perpendicular, then

$$(2\mathbf{a} + 3\mathbf{b}) \cdot (m\mathbf{a} + \mathbf{b}) = 0.$$

Expanding this expression,

$$2m\mathbf{a}\cdot\mathbf{a} + 2\mathbf{a}\cdot\mathbf{b} + 3m\mathbf{b}\cdot\mathbf{a} + 3\mathbf{b}\cdot\mathbf{b} = 0.$$

Now $\mathbf{a}$ and $\mathbf{b}$ are perpendicular, so $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a} = 0$, and they are unit vectors, so $\mathbf{a}\cdot\mathbf{a} = \mathbf{b}\cdot\mathbf{b} = 1$. Thus

$$2m + 3 = 0,$$

so $m = -1.5$.

(b) $$\begin{aligned}|\mathbf{c}|^2 &= \mathbf{c}\cdot\mathbf{c} \\ &= (3\mathbf{a} + 5\mathbf{b}) \cdot (3\mathbf{a} + 5\mathbf{b}) \\ &= 9\mathbf{a}\cdot\mathbf{a} + 15\mathbf{a}\cdot\mathbf{b} + 15\mathbf{b}\cdot\mathbf{a} + 25\mathbf{b}\cdot\mathbf{b}.\end{aligned}$$

Thus, since $\mathbf{a}$ and $\mathbf{b}$ are perpendicular unit vectors,

$$|\mathbf{c}|^2 = 9 + 25 = 34,$$

so $|\mathbf{c}| = \sqrt{34} \simeq 5.831$.

Solution to Exercise 17

$\mathbf{a}\cdot\mathbf{b} = (4 \times 1) + (1 \times -3) + (-5 \times 1) = -4.$

The negative sign tells us that the angle between $\mathbf{a}$ and $\mathbf{b}$ is between $\pi/2$ and π radians, that is, it is an obtuse angle.

Solution to Exercise 18

$|\mathbf{a}| = \sqrt{2^2 + (-3)^2 + 1^2} = \sqrt{14},$

$|\mathbf{b}| = \sqrt{(-1)^2 + 2^2 + 4^2} = \sqrt{21}.$

Also,

$$\mathbf{a}\cdot\mathbf{b} = (2 \times -1) + (-3 \times 2) + (1 \times 4) = -4,$$

so if θ is the angle between $\mathbf{a}$ and $\mathbf{b}$, then

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|} = \frac{-4}{\sqrt{14} \times \sqrt{21}} = -\frac{4}{7\sqrt{6}}.$$

The negative sign means that θ is obtuse; we have $\theta \simeq 1.806$ (radians).

Solution to Exercise 19

$$\begin{aligned}\mathbf{a}\cdot\mathbf{i} &= (a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k})\cdot\mathbf{i}\\ &= a_1\,\mathbf{i}\cdot\mathbf{i} + a_2\,\mathbf{j}\cdot\mathbf{i} + a_3\,\mathbf{k}\cdot\mathbf{i}\\ &= a_1.\end{aligned}$$

Similarly,

$$\mathbf{a}\cdot\mathbf{j} = a_2 \quad \text{and} \quad \mathbf{a}\cdot\mathbf{k} = a_3.$$

(Notice that this means that the components of a vector are given by the dot products of the vector with the Cartesian unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.)

Solution to Exercise 20

(a) $\mathbf{a}\cdot\mathbf{c} = -2$, $\mathbf{a}\cdot\mathbf{d} = -3$, $\mathbf{a}\cdot\mathbf{e} = 0$.

Thus only $\mathbf{e}$ is perpendicular to $\mathbf{a}$.

(b) First,

$$\mathbf{a} + 2\mathbf{b} = \mathbf{j} + 9\mathbf{k}.$$

Now a suitable vector along the line joining the origin to the point $(1, 1, 1)$ is $\mathbf{i} + \mathbf{j} + \mathbf{k}$. The corresponding unit vector is $\widehat{\mathbf{u}} = \frac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})$. The component of $\mathbf{a} + 2\mathbf{b}$ in the direction of this line is

$$\begin{aligned}(\mathbf{a} + 2\mathbf{b})\cdot\widehat{\mathbf{u}} &= (\mathbf{j} + 9\mathbf{k})\cdot\tfrac{1}{\sqrt{3}}(\mathbf{i} + \mathbf{j} + \mathbf{k})\\ &= \tfrac{10}{\sqrt{3}}.\end{aligned}$$

Solution to Exercise 21

The angle between $-\mathbf{i}$ and $\mathbf{p}$ is α, and the angle between $\mathbf{j}$ and $\mathbf{p}$ is $\frac{\pi}{2} - \alpha$. Therefore the $\mathbf{i}$-component of $\mathbf{p}$ is

$$-|\mathbf{p}|\cos(\alpha) = -2.5\cos\alpha,$$

and the $\mathbf{j}$-component of $\mathbf{p}$ is

$$|\mathbf{p}|\cos\left(\tfrac{\pi}{2} - \alpha\right) = 2.5\sin\alpha.$$

Hence

$$\mathbf{p} = -2.5\cos\alpha\,\mathbf{i} + 2.5\sin\alpha\,\mathbf{j}.$$

The angle between $-\mathbf{i}$ and $\mathbf{q}$ is β, and the angle between $-\mathbf{j}$ and $\mathbf{q}$ is $\frac{\pi}{2} - \beta$. Therefore the $\mathbf{i}$-component of $\mathbf{q}$ is

$$-|\mathbf{q}|\cos(\beta) = -3\cos\beta,$$

and the $\mathbf{j}$-component of $\mathbf{q}$ is

$$-|\mathbf{q}|\cos\left(\tfrac{\pi}{2} - \beta\right) = -3\sin\beta.$$

Hence

$$\mathbf{q} = -3\cos\beta\,\mathbf{i} - 3\sin\beta\,\mathbf{j}.$$

Finally, the angle between $\mathbf{i}$ and $\mathbf{r}$ is γ, and the angle between $-\mathbf{j}$ and $\mathbf{r}$ is $\frac{\pi}{2} - \gamma$. Therefore the $\mathbf{i}$-component of $\mathbf{r}$ is

$$|\mathbf{r}| \cos\gamma = 2.5 \cos\gamma,$$

and the $\mathbf{j}$-component of $\mathbf{r}$ is

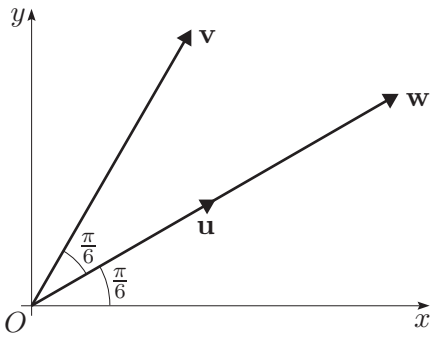
$$-|\mathbf{r}| \cos\left(\tfrac{\pi}{2} - \gamma\right) = -2.5 \sin\gamma.$$

Hence

$$\mathbf{r} = 2.5 \cos\gamma\, \mathbf{i} - 2.5 \sin\gamma\, \mathbf{j}.$$

Solution to Exercise 22

For the sake of clarity, here is a diagram showing $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ (where all three vectors start at O) drawn in the (x, y)-plane. (The z-axis points out of the page.)



The cross products are all perpendicular to the (x, y)-plane.

A unit vector in the direction of $\mathbf{u} \times \mathbf{v}$ is $\mathbf{k}$, so

$$\mathbf{u} \times \mathbf{v} = \left(|\mathbf{u}|\,|\mathbf{v}| \sin\tfrac{\pi}{6}\right) \mathbf{k} = \left(2 \times 3 \times \tfrac{1}{2}\right) \mathbf{k} = 3\mathbf{k}.$$

The angle between $\mathbf{u}$ and $\mathbf{w}$ is zero, so

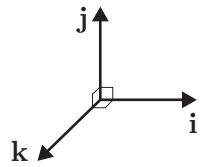
$$\mathbf{u} \times \mathbf{w} = \mathbf{0}.$$

A unit vector in the direction of $\mathbf{v} \times \mathbf{w}$ is $-\mathbf{k}$, so

$$\mathbf{v} \times \mathbf{w} = \left(|\mathbf{v}|\,|\mathbf{w}| \sin\tfrac{\pi}{6}\right) (-\mathbf{k}) = \left(3 \times 4 \times \tfrac{1}{2}\right) (-\mathbf{k}) = -6\mathbf{k}.$$

Solution to Exercise 23

(a) $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit vectors forming a right-handed system (see the figure in the margin).



Thus, using the definition of the cross product,

$$\mathbf{i} \times \mathbf{j} = \left(|\mathbf{i}|\,|\mathbf{j}| \sin\tfrac{\pi}{2}\right) \mathbf{k} = \mathbf{k}.$$

Similarly,

$$\mathbf{j} \times \mathbf{k} = \mathbf{i} \quad \text{and} \quad \mathbf{k} \times \mathbf{i} = \mathbf{j}.$$

(b) Since $(\mathbf{a} \times \mathbf{b}) = -(\mathbf{b} \times \mathbf{a})$ for any vectors $\mathbf{a}$ and $\mathbf{b}$, we have

$$\mathbf{j} \times \mathbf{i} = -\mathbf{k}, \quad \mathbf{k} \times \mathbf{j} = -\mathbf{i} \quad \text{and} \quad \mathbf{i} \times \mathbf{k} = -\mathbf{j}.$$

(c) Since $\mathbf{a} \times \mathbf{a} = \mathbf{0}$ for any vector $\mathbf{a}$, we have

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}.$$

(d)
$$\begin{aligned}(\mathbf{i}+\mathbf{k}) \times (\mathbf{i}+\mathbf{j}+\mathbf{k}) &= (\mathbf{i} \times (\mathbf{i}+\mathbf{j}+\mathbf{k})) + (\mathbf{k} \times (\mathbf{i}+\mathbf{j}+\mathbf{k})) \\ &= (\mathbf{0}+\mathbf{k}+(-\mathbf{j})) + (\mathbf{j}+(-\mathbf{i})+\mathbf{0}) \\ &= -\mathbf{i}+\mathbf{k},\end{aligned}$$

$$(\mathbf{i} \times (\mathbf{i}+\mathbf{k})) - ((\mathbf{i}+\mathbf{j}) \times \mathbf{k}) = (\mathbf{0}+(-\mathbf{j})) - (-\mathbf{j}+\mathbf{i}) = -\mathbf{i}.$$

(e)
$$\begin{aligned}(\mathbf{a}+\mathbf{b}) \times (\mathbf{a}+2\mathbf{b}) &= \mathbf{a} \times \mathbf{a} + \mathbf{a} \times 2\mathbf{b} + \mathbf{b} \times \mathbf{a} + \mathbf{b} \times 2\mathbf{b} \\ &= \mathbf{0} + 2\mathbf{a} \times \mathbf{b} - \mathbf{a} \times \mathbf{b} + \mathbf{0} \\ &= \mathbf{a} \times \mathbf{b},\end{aligned}$$

$$\begin{aligned}(\mathbf{a}+\mathbf{b}) \times (\mathbf{a}-\mathbf{b}) &= \mathbf{a} \times \mathbf{a} - \mathbf{a} \times \mathbf{b} + \mathbf{b} \times \mathbf{a} - \mathbf{b} \times \mathbf{b} \\ &= \mathbf{0} + \mathbf{a} \times \mathbf{b} - \mathbf{a} \times \mathbf{b} + \mathbf{0} = \mathbf{0}.\end{aligned}$$

Solution to Exercise 24

First, we list the components of the vectors:

$$a_1 = 2,\ a_2 = -3,\ a_3 = 1,$$
$$b_1 = -1,\ b_2 = 2,\ b_3 = 4,$$
$$c_1 = -4,\ c_2 = 6,\ c_3 = -2.$$

To compute $\mathbf{a} \times \mathbf{b}$, we need

$$a_2b_3 - a_3b_2 = -3 \times 4 - 1 \times 2 = -14,$$
$$a_3b_1 - a_1b_3 = 1 \times (-1) - 2 \times 4 = -9,$$
$$a_1b_2 - a_2b_1 = 2 \times 2 - (-3) \times (-1) = 1.$$

Hence $\mathbf{a} \times \mathbf{b} = -14\mathbf{i} - 9\mathbf{j} + \mathbf{k}$.

Similarly, for $\mathbf{a} \times \mathbf{c}$ we need

$$a_2c_3 - a_3c_2 = -3 \times (-2) - 1 \times 6 = 0,$$
$$a_3c_1 - a_1c_3 = 1 \times (-4) - 2 \times (-2) = 0,$$
$$a_1c_2 - a_2c_1 = 2 \times 6 - (-3) \times (-4) = 0.$$

Hence $\mathbf{a} \times \mathbf{c} = \mathbf{0}$.

Finally, for $\mathbf{b} \times \mathbf{c}$ we need

$$b_2c_3 - b_3c_2 = 2 \times (-2) - 4 \times 6 = -28,$$
$$b_3c_1 - b_1c_3 = 4 \times (-4) - (-1) \times (-2) = -18,$$
$$b_1c_2 - b_2c_1 = (-1) \times 6 - 2 \times (-4) = 2.$$

Hence $\mathbf{b} \times \mathbf{c} = -28\mathbf{i} - 18\mathbf{j} + 2\mathbf{k}$.

Since $\mathbf{a} \times \mathbf{c} = \mathbf{0}$, and neither vector is zero, the vectors $\mathbf{a}$ and $\mathbf{c}$ are parallel. In fact, $\mathbf{c} = -2\mathbf{a}$.

Solution to Exercise 25

A vector perpendicular to $\mathbf{a}$ and $\mathbf{b}$ is $\mathbf{a} \times \mathbf{b}$. The components of $\mathbf{a}$ and $\mathbf{b}$ are

$$a_1 = 2,\ a_2 = 2,\ a_3 = 1,$$
$$b_1 = 4,\ b_2 = 4,\ b_3 = -7.$$

So we have

$$a_2b_3 - a_3b_2 = 2 \times (-7) - 1 \times 4 = -18,$$
$$a_3b_1 - a_1b_3 = 1 \times 4 - 2 \times (-7) = 18,$$
$$a_1b_2 - a_2b_1 = 2 \times 4 - 2 \times 4 = 0.$$

Hence $\mathbf{a} \times \mathbf{b} = -18\mathbf{i} + 18\mathbf{j}$.

We are asked for a unit vector, so the obvious choice is

$$\frac{1}{|\mathbf{a} \times \mathbf{b}|}(\mathbf{a} \times \mathbf{b}) = \frac{-18\mathbf{i} + 18\mathbf{j}}{18\sqrt{2}} = \tfrac{1}{\sqrt{2}}(-\mathbf{i} + \mathbf{j}).$$

(Note that $\frac{1}{\sqrt{2}}(\mathbf{i} - \mathbf{j})$ is also a unit vector perpendicular to $\mathbf{a}$ and $\mathbf{b}$. This can be obtained by considering $\mathbf{b} \times \mathbf{a}$ rather than $\mathbf{a} \times \mathbf{b}$.)

Solution to Exercise 26

(a) $\mathbf{t}$ is perpendicular to both $\mathbf{r}$ and $\mathbf{s}$, and its sense is vertically down, that is, into the ground.

(b) Conversely, the sense of $\mathbf{s} \times \mathbf{r}$ is vertically up.

(c) $\mathbf{t} \times \mathbf{r}$ is perpendicular to $\mathbf{t}$ (thus in the horizontal plane) and perpendicular to $\mathbf{r}$, and by the right-hand grip rule its sense is due east.

(d) $|\mathbf{t}| = |\mathbf{r}|\,|\mathbf{s}| \sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}$.

(e) $\mathbf{t} \times (\mathbf{r} \times \mathbf{s}) = \mathbf{t} \times \mathbf{t} = \mathbf{0}$.

(f) $\mathbf{r} \cdot \mathbf{s} = |\mathbf{r}|\,|\mathbf{s}| \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$.

(g)
$$\begin{aligned}\mathbf{s} \cdot (\mathbf{t} \times \mathbf{r}) &= |\mathbf{s}|\,|\mathbf{t} \times \mathbf{r}| \cos\tfrac{\pi}{4} \quad \text{(by part (c))}\\ &= |\mathbf{s}|\left(|\mathbf{t}|\,|\mathbf{r}| \sin\tfrac{\pi}{2}\right)\cos\tfrac{\pi}{4}\\ &= 1 \times \tfrac{1}{\sqrt{2}} \times 1 \times 1 \times \tfrac{1}{\sqrt{2}} = \tfrac{1}{2}.\end{aligned}$$

Solution to Exercise 27

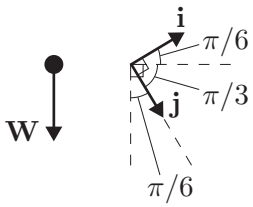
Two sides of the triangle are $\overrightarrow{AB} = -\mathbf{i} - \mathbf{j} + 5\mathbf{k}$ and $\overrightarrow{AC} = 2\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}$. So the area of the triangle is

$$\begin{aligned}\tfrac{1}{2}\left|\overrightarrow{AB} \times \overrightarrow{AC}\right| &= \tfrac{1}{2}\big|((-1) \times 2 - 5 \times (-3))\mathbf{i} + (5 \times 2 - (-1) \times 2)\mathbf{j}\\ &\qquad + ((-1) \times (-3) - (-1) \times 2)\mathbf{k}\big|\\ &= \tfrac{1}{2}|13\mathbf{i} + 12\mathbf{j} + 5\mathbf{k}|\\ &= \tfrac{1}{2}\sqrt{169 + 144 + 25}\\ &= \tfrac{13}{2}\sqrt{2}.\end{aligned}$$

Solution to Exercise 28

The weight is $3g\,\mathbf{k}$.

Solution to Exercise 29



Since the weight acts vertically downwards, its direction is in the vertical plane defined by $\mathbf{i}$ and $\mathbf{j}$, so $\mathbf{W}$ has no $\mathbf{k}$-component. To find the $\mathbf{i}$- and $\mathbf{j}$-components, the first step is to draw a diagram and work out the angles involved. In this case we use the two right angles marked in the diagram in the margin to work out the required angles.

A unit vector in the direction of $\mathbf{W}$ is $\widehat{\mathbf{u}} = -\sin\frac{\pi}{6}\,\mathbf{i} + \cos\frac{\pi}{6}\,\mathbf{j}$.

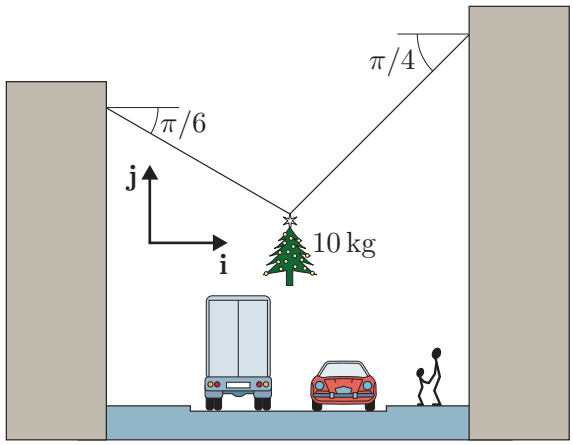
Resolving the weight gives

$$\begin{aligned}\mathbf{W} = |\mathbf{W}|\,\widehat{\mathbf{u}} &= \left(-|\mathbf{W}|\sin\tfrac{\pi}{6}\right)\mathbf{i} + \left(|\mathbf{W}|\cos\tfrac{\pi}{6}\right)\mathbf{j}\\ &= -15g \times -\tfrac{1}{2}\,\mathbf{i} + 15g \times \tfrac{\sqrt{3}}{2}\,\mathbf{j}\\ &= -\tfrac{15}{2}g\,\mathbf{i} + \tfrac{15\sqrt{3}}{2}g\,\mathbf{j}.\end{aligned}$$

Solution to Exercise 30

In this and other solutions, you may find that your diagrams and chosen axes are different from those given. You should still be able to check the validity of your solution against the given one, as the basic concepts are unchanged by these differences. Any choice of axes should lead to the same final answers as those given.

◀ Draw picture ▶

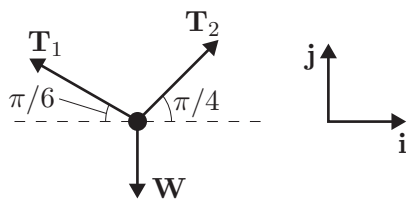


◀ Choose axes ▶

The forces all lie in a vertical plane, so we need only two axes. The unit vectors $\mathbf{i}$ and $\mathbf{j}$ are shown in the diagram.

◀ Draw force diagram(s) ▶

The force diagram, where the tension forces are denoted by $\mathbf{T}_1$ and $\mathbf{T}_2$, and the weight of the tree is denoted by $\mathbf{W}$, is as follows.



The equilibrium condition for particles gives ◀ Apply law(s) ▶

$$\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{W} = \mathbf{0}.$$

From the force diagram we have ◀ Solve equation(s) ▶

$$\mathbf{W} = -|\mathbf{W}|\,\mathbf{j} = -10g\,\mathbf{j}.$$

The other forces can be expressed in terms of components:

$$\mathbf{T}_1 = -|\mathbf{T}_1|\cos\tfrac{\pi}{6}\,\mathbf{i} + |\mathbf{T}_1|\sin\tfrac{\pi}{6}\,\mathbf{j} = -\tfrac{\sqrt{3}}{2}|\mathbf{T}_1|\,\mathbf{i} + \tfrac{1}{2}|\mathbf{T}_1|\,\mathbf{j},$$
$$\mathbf{T}_2 = |\mathbf{T}_2|\cos\tfrac{\pi}{4}\,\mathbf{i} + |\mathbf{T}_2|\sin\tfrac{\pi}{4}\,\mathbf{j} = \tfrac{1}{\sqrt{2}}|\mathbf{T}_2|\,\mathbf{i} + \tfrac{1}{\sqrt{2}}|\mathbf{T}_2|\,\mathbf{j}.$$

Resolving the equilibrium equation in the **i**-direction gives

$$-\tfrac{\sqrt{3}}{2}|\mathbf{T}_1| + \tfrac{1}{\sqrt{2}}|\mathbf{T}_2| + 0 = 0,$$

and resolving in the **j**-direction gives

$$\tfrac{1}{2}|\mathbf{T}_1| + \tfrac{1}{\sqrt{2}}|\mathbf{T}_2| - 10g = 0.$$

Subtracting these two equations gives

$$\left(\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}\right)|\mathbf{T}_1| - 10g = 0,$$

so

$$|\mathbf{T}_1| = \frac{20}{1+\sqrt{3}}\,g \simeq 71.81.$$

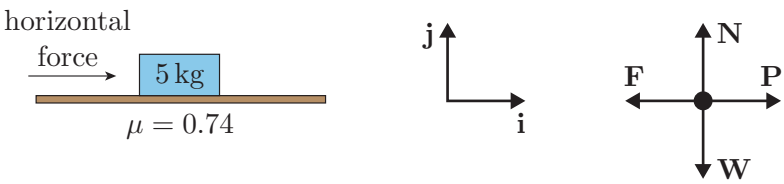
Substituting this value of $|\mathbf{T}_1|$ into the **i**-direction equation gives

$$|\mathbf{T}_2| = \frac{20\sqrt{3}}{\sqrt{2}+\sqrt{6}}\,g \simeq 87.95.$$

The model predicts that the magnitudes of the tension forces due to the ropes are 72 N and 88 N to the nearest whole number. ◀ Interpret solution ▶

Solution to Exercise 31

Draw the picture, choose axes and draw the force diagram. ◀ Draw picture ▶ ◀ Choose axes ▶ ◀ Draw force diagram(s) ▶



If the object does not move, the equilibrium condition for particles holds, so ◀ Apply law(s) ▶

$$\mathbf{F} + \mathbf{N} + \mathbf{P} + \mathbf{W} = \mathbf{0}.$$

If the block does not move,

$$|\mathbf{F}| \le \mu|\mathbf{N}|,$$

where μ is the coefficient of static friction.

◀ Solve equation(s) ▶

Resolving the equilibrium equation in the **i**-direction gives

$$|\mathbf{F}| = |\mathbf{P}|.$$

Resolving in the **j**-direction gives

$$|\mathbf{N}| = |\mathbf{W}|.$$

Therefore if the block does not move,

$$|\mathbf{P}| = |\mathbf{F}| \leq \mu|\mathbf{N}| = \mu|\mathbf{W}|.$$

Since $|\mathbf{P}| = 2$ and $|\mathbf{W}| = 0.5g$, the block does not move provided that

$$\mu \geq \frac{2}{0.5g} \simeq 0.41.$$

◀ Interpret solution ▶

The coefficient of static friction for steel on dry steel is $\mu = 0.74$, so the model predicts that the block will not move.

Since we have $|\mathbf{F}| = |\mathbf{P}|$, the magnitude of the friction force is 2 N.

Solution to Exercise 32

Since the friction force does not depend on the area of contact, the sideways force required to start the box slipping is the same in both cases.

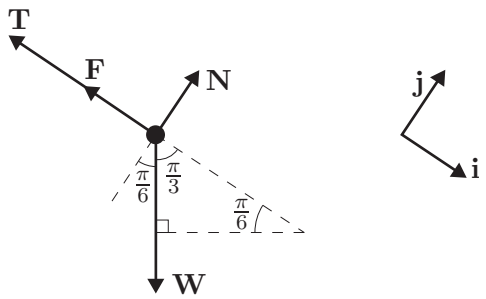
Solution to Exercise 33

◀ Choose axes ▶

Choose the same axes as in Example 9.

◀ Draw force diagram(s) ▶

Modelling the crate as a particle and the rope as a model string, the force diagram is as shown below, where the notation is as in Example 9. Both **F** and **T** act in the same direction (on the implied assumption that the crate will slide down the ramp if nothing holds it back).



(It is conventional to draw two forces acting in the same direction on a particle as overlapping arrows of different lengths.)

◀ Apply law(s) ▶

The equilibrium condition for particles gives

$$\mathbf{F} + \mathbf{T} + \mathbf{N} + \mathbf{W} = \mathbf{0}.$$

If the crate does not move, then

$$|\mathbf{F}| \leq \mu|\mathbf{N}| = 0.2|\mathbf{N}|,$$

where μ is the coefficient of static friction.

From the diagram, we have ◀ Solve equations ▶

$$\mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{T} = |\mathbf{T}|\,(-\mathbf{i}) \quad \text{and} \quad \mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}).$$

Resolving gives

$$\mathbf{W} = |\mathbf{W}|\cos\tfrac{\pi}{3}\,\mathbf{i} - |\mathbf{W}|\sin\tfrac{\pi}{3}\,\mathbf{j} = \tfrac{1}{2}|\mathbf{W}|\,\mathbf{i} - \tfrac{\sqrt{3}}{2}|\mathbf{W}|\,\mathbf{j}.$$

Resolving the equilibrium equation in the **i**-direction gives

$$-|\mathbf{F}| - |\mathbf{T}| + 0 + \tfrac{1}{2}|\mathbf{W}| = 0,$$

so

$$|\mathbf{F}| = \tfrac{1}{2}|\mathbf{W}| - |\mathbf{T}|.$$

Resolving in the **j**-direction gives

$$0 + 0 + |\mathbf{N}| - \tfrac{\sqrt{3}}{2}|\mathbf{W}| = 0,$$

so

$$|\mathbf{N}| = \tfrac{\sqrt{3}}{2}|\mathbf{W}|.$$

Substituting for $|\mathbf{F}|$ and $|\mathbf{N}|$ into the friction condition gives

$$\tfrac{1}{2}|\mathbf{W}| - |\mathbf{T}| \leq 0.2 \times \tfrac{\sqrt{3}}{2}|\mathbf{W}|,$$

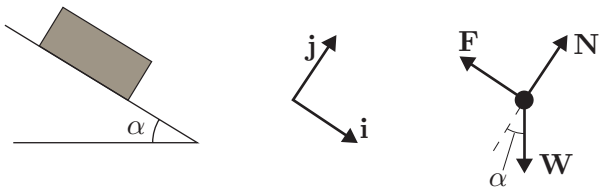
which, on rearrangement and using $|\mathbf{W}| = 60g$, gives

$$|\mathbf{T}| \geq 30g - 6\sqrt{3}g \simeq 192.$$

Therefore the model predicts that a force of at least 192 N up the slope needs to be applied to the rope to keep the crate from sliding down the ramp. ◀ Interpret solution ▶

Solution to Exercise 34

(a) The situation is illustrated below. ◀ Draw picture ▶ ◀ Draw force diagram(s) ▶



Choose axes so that **i** points down the slope and **j** is in the direction of the normal reaction. ◀ Choose axes ▶

The equilibrium condition for particles gives ◀ Apply law(s) ▶

$$\mathbf{F} + \mathbf{N} + \mathbf{W} = \mathbf{0}.$$

Since the box is on the point of slipping,

$$|\mathbf{F}| = \mu|\mathbf{N}|,$$

where μ is the coefficient of static friction.

◀ Solve equation(s) ▶

Two of the forces are aligned with the axes and can be written down immediately:

$$\mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}) \quad \text{and} \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}.$$

The third force is inclined to the axes, so resolving gives

$$\begin{aligned}\mathbf{W} &= |\mathbf{W}|\cos\left(\tfrac{\pi}{2} - \alpha\right)\mathbf{i} - |\mathbf{W}|\sin\left(\tfrac{\pi}{2} - \alpha\right)\mathbf{j} \\ &= mg\sin\alpha\,\mathbf{i} - mg\cos\alpha\,\mathbf{j}.\end{aligned}$$

Resolving the equilibrium equation in the **i**-direction gives

$$-|\mathbf{F}| + 0 + mg\sin\alpha = 0,$$

so

$$|\mathbf{F}| = mg\sin\alpha.$$

Resolving in the **j**-direction gives

$$0 + |\mathbf{N}| - mg\cos\alpha = 0,$$

so

$$|\mathbf{N}| = mg\cos\alpha.$$

Substituting $|\mathbf{F}|$ and $|\mathbf{N}|$ into the friction condition gives

$$\mu = \frac{|\mathbf{F}|}{|\mathbf{N}|} = \tan\alpha.$$

◀ Interpret solution ▶

(b) The result in part (a) tells us that for an object on an inclined plane, the angle at which the object starts to slip depends only on the two surfaces in contact (i.e. on the coefficient of static friction μ). The mass of the object is irrelevant – so the half-full mug will start to slip at the same angle as the empty one.

Solution to Exercise 35

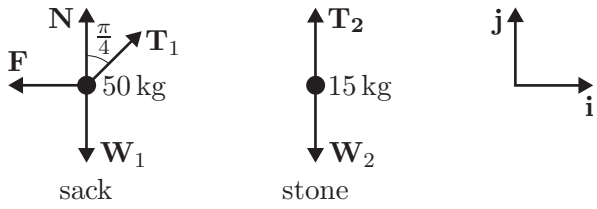
(a) The sack, stone, pulley and rope are modelled as in Example 10.

◀ Choose axes ▶

This time we need two axes, as shown in Figure 55.

◀ Draw force diagram(s) ▶

The force diagrams for the sack and the stone, using the usual notation, are as follows.



◀ Apply law(s) ▶

The equilibrium condition for particles gives

$$\mathbf{W}_1 + \mathbf{F} + \mathbf{N} + \mathbf{T}_1 = \mathbf{0},$$
$$\mathbf{W}_2 + \mathbf{T}_2 = \mathbf{0}.$$

The fact that we have a model pulley tells us that

$$|\mathbf{T}_1| = |\mathbf{T}_2|.$$

From the diagram we have ◀ Solve equation(s) ▶

$$\mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}), \quad \mathbf{W}_1 = |\mathbf{W}_1|\,(-\mathbf{j}),$$
$$\mathbf{T}_2 = |\mathbf{T}_2|\,\mathbf{j}, \quad \mathbf{W}_2 = |\mathbf{W}_2|\,(-\mathbf{j}).$$

Resolving for $\mathbf{T}_1$ gives

$$\mathbf{T}_1 = |\mathbf{T}_1| \cos \tfrac{\pi}{4}\,\mathbf{i} + |\mathbf{T}_1| \sin \tfrac{\pi}{4}\,\mathbf{j} = \tfrac{1}{\sqrt{2}}|\mathbf{T}_1|\,\mathbf{i} + \tfrac{1}{\sqrt{2}}|\mathbf{T}_1|\,\mathbf{j}.$$

Resolving the first equilibrium equation in the **i**-direction gives

$$0 - |\mathbf{F}| + 0 + \tfrac{1}{\sqrt{2}}|\mathbf{T}_1| = 0,$$

so

$$|\mathbf{F}| = \tfrac{1}{\sqrt{2}}|\mathbf{T}_1|.$$

Resolving both equilibrium equations in the **j**-direction gives

$$-|\mathbf{W}_1| + 0 + |\mathbf{N}| + \tfrac{1}{\sqrt{2}}|\mathbf{T}_1| = 0,$$
$$-|\mathbf{W}_2| + |\mathbf{T}_2| = 0,$$

so

$$|\mathbf{N}| = |\mathbf{W}_1| - \tfrac{1}{\sqrt{2}}|\mathbf{T}_1| \quad \text{and} \quad |\mathbf{W}_2| = |\mathbf{T}_2|.$$

(b) We have ◀ Interpret solution ▶

$$|\mathbf{W}_1| = 50g, \quad |\mathbf{W}_2| = 15g, \quad |\mathbf{T}_1| = |\mathbf{T}_2| = |\mathbf{W}_2|,$$

so

$$|\mathbf{N}| = |\mathbf{W}_1| - \tfrac{1}{\sqrt{2}}|\mathbf{T}_1| = 50g - \tfrac{15}{\sqrt{2}}g \simeq 386.$$

So the model predicts that the magnitude of the normal reaction is about 386 N.

(c) We have

$$|\mathbf{F}| = \tfrac{1}{\sqrt{2}}|\mathbf{T}_1| = \tfrac{1}{\sqrt{2}}|\mathbf{W}_2| = \tfrac{15}{\sqrt{2}}g \simeq 104,$$

so the model predicts that the magnitude of the friction force is about 104 N.

(d) For equilibrium, we must have $|\mathbf{F}| \leq \mu|\mathbf{N}|$, that is,

$$\mu \geq |\mathbf{F}|/|\mathbf{N}| \simeq 104/386 \simeq 0.27.$$

So the model predicts that the smallest value of μ that allows the system to remain in equilibrium is about 0.27.

(e) If $|\mathbf{N}|$ approaches zero, then $|\mathbf{W}_1| - \frac{1}{\sqrt{2}}|\mathbf{T}_1| = |\mathbf{W}_1| - \frac{1}{\sqrt{2}}|\mathbf{W}_2|$ approaches zero, so $|\mathbf{W}_2|$ tends to $|\sqrt{2}\,\mathbf{W}_1|$. Thus the stone would have to weigh approximately 70.7 kg.

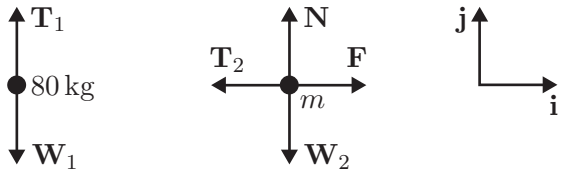
Solution to Exercise 36

Model the man in the well as a particle of mass 80 kg, and all the men on the ground as a single particle of mass $m = 80n$, where n is the number of men.

◀ Choose axes ▶

◀ Draw force diagram(s) ▶

A suitable choice of axes is shown in Figure 57. A force diagram for each particle, using the usual notation, is shown below.



man in well men on ground

◀ Apply law(s) ▶

The equilibrium condition for particles gives

$$\mathbf{T}_1 + \mathbf{W}_1 = \mathbf{0},$$
$$\mathbf{T}_2 + \mathbf{N} + \mathbf{F} + \mathbf{W}_2 = \mathbf{0}.$$

The assumption of a model pulley gives

$$|\mathbf{T}_1| = |\mathbf{T}_2|.$$

If the men are not to slip, we must have

$$|\mathbf{F}| \leq \mu|\mathbf{N}|,$$

where μ is the coefficient of static friction.

◀ Solve equation(s) ▶

Resolving the second equilibrium equation in the **i**- and **j**-directions in turn gives

$$-|\mathbf{T}_2| + 0 + |\mathbf{F}| + 0 = 0,$$
$$0 + |\mathbf{N}| + 0 - |\mathbf{W}_2| = 0.$$

Resolving the first equilibrium equation in the **j**-direction gives

$$|\mathbf{T}_1| - |\mathbf{W}_1| = 0.$$

Therefore

$$|\mathbf{F}| = |\mathbf{T}_2| = |\mathbf{T}_1| = |\mathbf{W}_1| = 80g,$$
$$|\mathbf{N}| = |\mathbf{W}_2| = mg = 80ng,$$

hence

$$\mu|\mathbf{N}| = 0.45 \times mg = 0.45mg.$$

So, for the friction condition to be satisfied, we need $0.45 \times 80ng \geq 80g$, which gives $n \geq 1/0.45$. As n is an integer, the minimum value of n is 3.

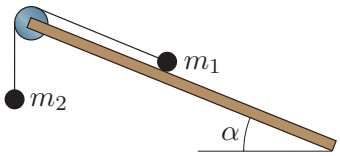
◀ Interpret solution ▶

The model predicts that three men each of mass 80 kg are required to hold the man above the well.

Solution to Exercise 37

Equilibrium can be disturbed either by the object of mass m_1 sliding down the board and pulling up the object of mass m_2, or by the object of mass m_2 dropping down and pulling the object of mass m_1 up the board.

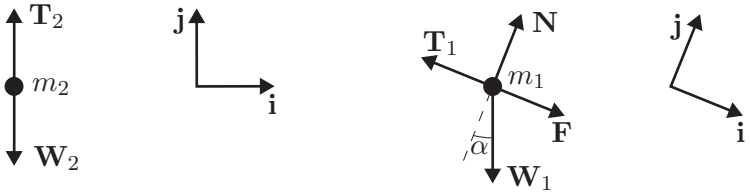
Model the masses as particles joined by a model string hanging over a model pulley representing the edge of the board. ◀ Draw picture ▶



We need to consider the two cases of possible movement separately, as they lead to different force diagrams and hence to different results.

In this problem it makes sense to use different sets of unit vectors for the two particles, as indicated below. ◀ Choose axes ▶

Consider first the case where m_2 is likely to drop down and pull m_1 up the board. The force diagrams are as follows. ◀ Draw force diagram(s) ▶



The equilibrium condition for particles gives ◀ Apply law(s) ▶

$$\mathbf{T}_1 + \mathbf{N} + \mathbf{F} + \mathbf{W}_1 = \mathbf{0},$$
$$\mathbf{T}_2 + \mathbf{W}_2 = \mathbf{0}.$$

The use of a model pulley gives

$$|\mathbf{T}_1| = |\mathbf{T}_2|.$$

If the system is to remain in equilibrium, we require

$$|\mathbf{F}| \leq \mu |\mathbf{N}|.$$

For the particle of mass m_2, the forces are $\mathbf{T}_2 = |\mathbf{T}_2|\,\mathbf{j}$ and $\mathbf{W}_2 = -|\mathbf{W}_2|\,\mathbf{j}$. From the second equilibrium equation we have $|\mathbf{T}_2|\,\mathbf{j} - |\mathbf{W}_2|\,\mathbf{j} = \mathbf{0}$, therefore ◀ Solve equation(s) ▶

$$|\mathbf{T}_2| = |\mathbf{W}_2| = m_2 g.$$

So by the model pulley equation,

$$|\mathbf{T}_1| = m_2 g.$$

For the particle of mass m_1, the forces acting are

$$\mathbf{F} = |\mathbf{F}|\,\mathbf{i}, \quad \mathbf{N} = |\mathbf{N}|\,\mathbf{j}, \quad \mathbf{T}_1 = -|\mathbf{T}_1|\,\mathbf{i} = -m_2 g\,\mathbf{i}$$

and $\mathbf{W}_1$, which can be resolved as

$$\mathbf{W}_1 = |\mathbf{W}_1| \sin\alpha\, \mathbf{i} + |\mathbf{W}_1| \cos\alpha\, (-\mathbf{j}).$$

Resolving the first equilibrium equation in the **i**- and **j**-directions in turn gives

$$-|\mathbf{T}_1| + 0 + |\mathbf{F}| + |\mathbf{W}_1| \sin\alpha = 0,$$
$$0 + |\mathbf{N}| + 0 - |\mathbf{W}_1| \cos\alpha = 0.$$

Therefore

$$|\mathbf{F}| = |\mathbf{T}_1| - |\mathbf{W}_1| \sin\alpha = m_2 g - m_1 g \sin\alpha,$$
$$|\mathbf{N}| = |\mathbf{W}_1| \cos\alpha = m_1 g \cos\alpha.$$

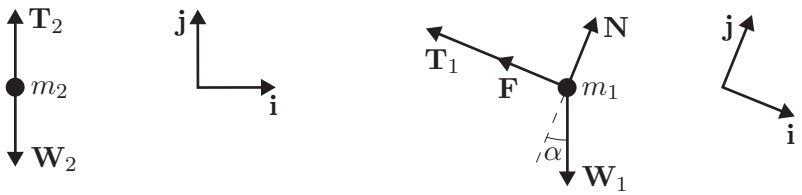
◀ Interpret solution ▶

For the friction condition to be satisfied, the model predicts that

$$\mu \geq \frac{|\mathbf{F}|}{|\mathbf{N}|} = \frac{m_2 g - m_1 g \sin\alpha}{m_1 g \cos\alpha} = \frac{m_2 - m_1 \sin\alpha}{m_1 \cos\alpha}.$$

◀ Draw force diagram(s) ▶

Now consider the other case where equilibrium could be disturbed, that is, where m_1 may slide down the board and pull m_2 up. The force diagrams are as follows.



◀ Apply law(s) ▶

The equilibrium condition and the use of a model pulley lead to the same equilibrium equations as before.

◀ Solve equation(s) ▶

The only difference is that the friction force is in the opposite direction, that is, $\mathbf{F} = -|\mathbf{F}|\,\mathbf{i}$. This means that when the first equilibrium equation is resolved in the **i**-direction, it now gives

$$-|\mathbf{T}_1| + 0 - |\mathbf{F}| + |\mathbf{W}_1| \sin\alpha = 0,$$

leading to

$$|\mathbf{F}| = |\mathbf{W}_1| \sin\alpha - |\mathbf{T}_1| = m_1 g \sin\alpha - m_2 g.$$

◀ Interpret solution ▶

For the friction condition to be satisfied, the model predicts that

$$\mu \geq \frac{|\mathbf{F}|}{|\mathbf{N}|} = \frac{m_1 \sin\alpha - m_2}{m_1 \cos\alpha}.$$

Both conditions for μ must hold for the system to be in equilibrium. Since μ is non-negative and the right-hand sides are of opposite signs, we can say

the first condition applies when $m_2 > m_1 \sin\alpha$,
the second condition applies when $m_2 < m_1 \sin\alpha$.

If there is no friction, then $\mu = 0$, and the conditions can both be satisfied only if the right-hand sides are also zero, that is, $m_2 = m_1 \sin\alpha$.

Solution to Exercise 38

We can model the seesaw as a model rod. Let Jill's mass be m. Then, assuming that neither Jack's nor Jill's feet are on the ground, we can use equation (31) to obtain

$$mg \times 1.2 = 60g \times 1.$$

Therefore $m = 50$, so Jill's mass is 50 kg.

Solution to Exercise 39

The simplest choice of position vector for a point on the line of action of $\mathbf{F}_1$ is $\mathbf{r}_1 = 2\mathbf{i}$, the position vector of the point $(2, 0, 0)$ relative to the origin. Similarly, suitably simple choices for $\mathbf{F}_2$ and $\mathbf{F}_3$ are $\mathbf{r}_2 = -2\mathbf{i}$ and $\mathbf{r}_3 = 2\mathbf{k}$. Hence the torques of the three forces relative to the origin are

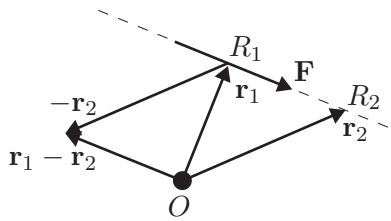
$$\mathbf{r}_1 \times \mathbf{F}_1 = 2\mathbf{i} \times 3\mathbf{j} = 6\mathbf{k},$$
$$\mathbf{r}_2 \times \mathbf{F}_2 = -2\mathbf{i} \times 3\mathbf{j} = -6\mathbf{k},$$
$$\mathbf{r}_3 \times \mathbf{F}_3 = 2\mathbf{k} \times 3\mathbf{j} = -6\mathbf{i}.$$

Solution to Exercise 40

(a) If O is on the line of action of $\mathbf{F}$, and $\mathbf{r}$ is the position vector relative to O of another point on that line of action, then $\mathbf{r}$ is parallel to $\mathbf{F}$, so $\mathbf{r} \times \mathbf{F} = \mathbf{0}$.

Thus when O is on the line of action of a force $\mathbf{F}$, the formula $\boldsymbol{\Gamma} = \mathbf{r} \times \mathbf{F}$ gives $\mathbf{0}$ whatever the choice of $\mathbf{r}$.

(b) If $\mathbf{r}_1$ and $\mathbf{r}_2$ are position vectors, relative to O, of two points R_1 and R_2 on the line of action of $\mathbf{F}$ (where O is not on this line of action), then the triangle rule for the addition of vectors shows that $\mathbf{r}_1 - \mathbf{r}_2$ is parallel to $\mathbf{F}$ (as shown in the diagram in the margin).



Therefore $(\mathbf{r}_1 - \mathbf{r}_2) \times \mathbf{F} = \mathbf{0}$, and since

$$(\mathbf{r}_1 - \mathbf{r}_2) \times \mathbf{F} = \mathbf{r}_1 \times \mathbf{F} - \mathbf{r}_2 \times \mathbf{F},$$

we have

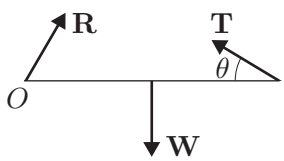
$$\mathbf{r}_1 \times \mathbf{F} = \mathbf{r}_2 \times \mathbf{F}.$$

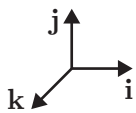
Thus the formula $\boldsymbol{\Gamma} = \mathbf{r} \times \mathbf{F}$ gives the same vector, whatever the choice of $\mathbf{r}$.

Solution to Exercise 41

◀ Draw force diagram(s) ▶

There are three forces acting on the rod: the weight $\mathbf{W}$ (at the centre of the rod), the tension force $\mathbf{T}$ due to the string, and the reaction force $\mathbf{R}$ at the hinge. The direction of $\mathbf{R}$ is not known, except that it is in the plane of $\mathbf{W}$ and $\mathbf{T}$. The force diagram is shown below.





◀ Choose axes ▶

Choose the origin to be at O and axes as shown above.

◀ Apply law(s) ▶

The equilibrium condition for rigid bodies gives

$$\mathbf{R} + \mathbf{T} + \mathbf{W} = \mathbf{0},$$
$$\boldsymbol{\Gamma}_{\mathrm{R}} + \boldsymbol{\Gamma}_{\mathrm{T}} + \boldsymbol{\Gamma}_{\mathrm{W}} = \mathbf{0}.$$

◀ Solve equation(s) ▶

Position vectors of the points of application are

$$\mathbf{r}_{\mathrm{R}} = \mathbf{0}, \quad \mathbf{r}_{\mathrm{T}} = l\,\mathbf{i}, \quad \mathbf{r}_{\mathrm{W}} = \tfrac{1}{2}l\,\mathbf{i}.$$

From the force diagram,

$$\mathbf{W} = mg(-\mathbf{j}).$$

The force $\mathbf{R}$ is unknown except that it is in the plane of $\mathbf{W}$ and $\mathbf{T}$, that is, the plane defined by $\mathbf{i}$ and $\mathbf{j}$. For convenience we write $\mathbf{R}$ in component form as

$$\mathbf{R} = R_i\,\mathbf{i} + R_j\,\mathbf{j},$$

where R_i and R_j are the $\mathbf{i}$- and $\mathbf{j}$-components of $\mathbf{R}$.

The remaining force can be resolved as

$$\mathbf{T} = |\mathbf{T}|\cos\theta\,(-\mathbf{i}) + |\mathbf{T}|\sin\theta\,\mathbf{j} = -|\mathbf{T}|\cos\theta\,\mathbf{i} + |\mathbf{T}|\sin\theta\,\mathbf{j}.$$

The torques are given by

$$\boldsymbol{\Gamma}_{\mathrm{R}} = \mathbf{0},$$
$$\boldsymbol{\Gamma}_{\mathrm{T}} = \mathbf{r}_{\mathrm{T}} \times \mathbf{T} = l\,\mathbf{i} \times |\mathbf{T}|(-\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) = l\,|\mathbf{T}|\sin\theta\,\mathbf{k},$$
$$\boldsymbol{\Gamma}_{\mathrm{W}} = \mathbf{r}_{\mathrm{W}} \times \mathbf{W} = \tfrac{1}{2}l\,\mathbf{i} \times mg(-\mathbf{j}) = -\tfrac{1}{2}lmg\,\mathbf{k}.$$

Resolving the second equilibrium equation in the $\mathbf{k}$-direction gives

$$l\,|\mathbf{T}|\sin\theta - \tfrac{1}{2}lmg = 0,$$

so

$$|\mathbf{T}| = \tfrac{1}{2}mg\,\mathrm{cosec}\,\theta.$$

Resolving the first equilibrium equation in the $\mathbf{i}$- and $\mathbf{j}$-directions in turn gives

$$R_i - |\mathbf{T}|\cos\theta = 0,$$
$$R_j + |\mathbf{T}|\sin\theta - mg = 0,$$

so

$$R_i = |\mathbf{T}|\cos\theta = \tfrac{1}{2}mg\cot\theta,$$
$$R_j = mg - |\mathbf{T}|\sin\theta = mg - \tfrac{1}{2}mg = \tfrac{1}{2}mg.$$

So the model predicts that the reaction force at the hinge is

$$\mathbf{R} = \tfrac{1}{2}mg\cot\theta\,\mathbf{i} + \tfrac{1}{2}mg\,\mathbf{j}.$$

From the first equilibrium equation, the tension force due to the string is

$$\mathbf{T} = -\mathbf{R} - \mathbf{W} = -\tfrac{1}{2}mg\cot\theta\,\mathbf{i} + \tfrac{1}{2}mg\,\mathbf{j}.$$

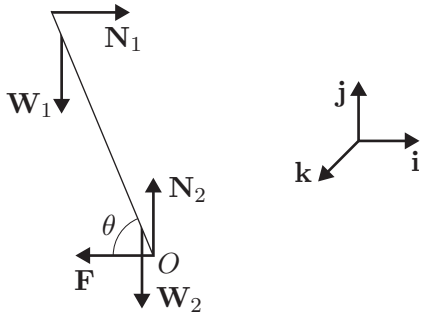
The **j**-components of **R** and **T** are the same. The **i**-components are of equal magnitude, but with opposite signs. Thus **R** and **T** have equal magnitude and make the same angle θ with the horizontal, where θ is measured anticlockwise for **R** and clockwise for **T**.

◀ Interpret solution ▶

Solution to Exercise 42

Since the mass of the ladder may be neglected, the forces acting on the ladder are the weights $\mathbf{W}_1$ and $\mathbf{W}_2$ of the two people, the normal reaction forces $\mathbf{N}_1$ and $\mathbf{N}_2$ at the wall and at the floor, and the friction force $\mathbf{F}$ at the floor. (There is no friction force at the wall, as it is smooth.) The force diagram is shown below.

◀ Draw force diagram(s) ▶



The origin O is best chosen to be at the foot of the ladder. The axes are shown above.

◀ Choose axes ▶

The equilibrium condition for rigid bodies gives

◀ Apply law(s) ▶

$$\mathbf{N}_1 + \mathbf{N}_2 + \mathbf{W}_1 + \mathbf{W}_2 + \mathbf{F} = \mathbf{0},$$
$$\boldsymbol{\Gamma}_{\mathrm{N}_1} + \boldsymbol{\Gamma}_{\mathrm{N}_2} + \boldsymbol{\Gamma}_{\mathrm{W}_1} + \boldsymbol{\Gamma}_{\mathrm{W}_2} + \boldsymbol{\Gamma}_{\mathrm{F}} = \mathbf{0}.$$

At the point of slipping,

$$|\mathbf{F}| = \mu |\mathbf{N}_2|,$$

where $\mu = 0.25$ is the coefficient of static friction.

Let l denote the length of the ladder (i.e. 3.9 m), and let d denote the distance of the end rungs from the ends of the ladder (i.e. 0.3 m). Relative to the origin O at the bottom of the ladder, two position vectors are zero, namely $\mathbf{r}_{\mathrm{N}_2}$ and $\mathbf{r}_{\mathrm{F}}$. The remaining position vectors are

◀ Solve equation(s) ▶

$$\mathbf{r}_{\mathrm{N}_1} = -l\cos\theta\,\mathbf{i} + l\sin\theta\,\mathbf{j},$$
$$\mathbf{r}_{\mathrm{W}_1} = -(l-d)\cos\theta\,\mathbf{i} + (l-d)\sin\theta\,\mathbf{j},$$
$$\mathbf{r}_{\mathrm{W}_2} = -d\cos\theta\,\mathbf{i} + d\sin\theta\,\mathbf{j}.$$

The forces are aligned with the axes, and can be put into component form by inspection:

$$\mathbf{N}_1 = |\mathbf{N}_1|\,\mathbf{i}, \quad \mathbf{N}_2 = |\mathbf{N}_2|\,\mathbf{j},$$
$$\mathbf{W}_1 = Mg\,(-\mathbf{j}), \quad \mathbf{W}_2 = mg\,(-\mathbf{j}), \quad \mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}),$$

where m is the mass of the person on the lower rung, and M is the mass of the person on the upper rung (i.e. 100 kg).

Two of the torques are zero by choice of origin. The remaining three torques can be calculated as

$$\boldsymbol{\Gamma}_{\mathrm{N}_1} = (-l\cos\theta\,\mathbf{i} + l\sin\theta\,\mathbf{j}) \times (|\mathbf{N}_1|\,\mathbf{i}) = -l\,|\mathbf{N}_1|\sin\theta\,\mathbf{k},$$
$$\boldsymbol{\Gamma}_{\mathrm{W}_1} = (-(l-d)\cos\theta\,\mathbf{i} + (l-d)\sin\theta\,\mathbf{j}) \times (-Mg\,\mathbf{j}) = M(l-d)g\cos\theta\,\mathbf{k},$$
$$\boldsymbol{\Gamma}_{\mathrm{W}_2} = (-d\cos\theta\,\mathbf{i} + d\sin\theta\,\mathbf{j}) \times (-mg\,\mathbf{j}) = mdg\cos\theta\,\mathbf{k}.$$

Substituting these into the second equilibrium equation and resolving in the **k**-direction gives

$$-l\,|\mathbf{N}_1|\sin\theta + M(l-d)g\cos\theta + mdg\cos\theta = 0.$$

Now we need to find the magnitude of the normal reaction $|\mathbf{N}_1|$, which we do by resolving the first equilibrium equation in the **i**- and **j**-directions in turn, to obtain

$$|\mathbf{N}_1| - |\mathbf{F}| = 0,$$
$$|\mathbf{N}_2| - Mg - mg = 0.$$

From these equations and the friction condition, we have

$$|\mathbf{N}_1| = |\mathbf{F}| = \mu|\mathbf{N}_2| = \mu(Mg + mg).$$

Substituting for $|\mathbf{N}_1|$ in the **k**-direction equation gives

$$-l\mu(Mg + mg)\sin\theta + M(l-d)g\cos\theta + mdg\cos\theta = 0,$$

thus

$$-Ml\mu\tan\theta - ml\mu\tan\theta + M(l-d) + md = 0,$$

so

$$m = M\,\frac{l\mu\tan\theta - l + d}{d - l\mu\tan\theta}.$$

◀ Interpret solution ▶

To use this formula, we need the value of $\tan\theta$, which by trigonometry is equal to $\sqrt{3.9^2 - 1.5^2}/1.5 = 2.4$. Now we substitute in the known values, namely $\mu = 0.25$, $d = 0.3$, $l = 3.9$ and $M = 100$, to find

$$m = 100 \times \frac{3.9 \times 0.25 \times 2.4 - 3.9 + 0.3}{0.3 - 3.9 \times 0.25 \times 2.4} \simeq 61.8.$$

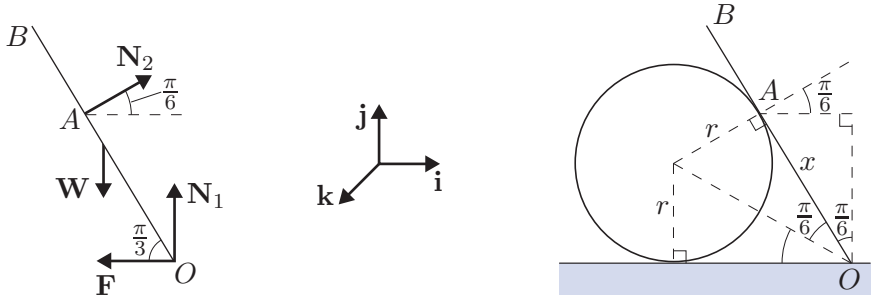
So to prevent the ladder from slipping, the model predicts that the person on the bottom rung must have a mass of at least 61.8 kg.

Solution to Exercise 43

◀ Choose axes ▶

◀ Draw force diagram(s) ▶

Take the origin at O at the base of the plank, and axes as shown below. The force diagram, using the usual notation, is also shown below.



To calculate the torque of the force $\mathbf{N}_2$ about O, we need the angle between $\mathbf{N}_2$ and $\mathbf{i}$, which from the diagram is equal to $\pi/6$. We also need the distance OA along the plank to the point of contact of the pipe and the plank: call this distance x. From the diagram, we have $\tan\frac{\pi}{6} = r/x$, which gives $x = \sqrt{3}r$. Now that we know the distances and angles involved, we can continue with Procedure 2.

The equilibrium condition for rigid bodies gives ◀ Apply law(s) ▶

$$\mathbf{N}_1 + \mathbf{N}_2 + \mathbf{F} + \mathbf{W} = \mathbf{0},$$
$$\boldsymbol{\Gamma}_{\mathrm{N}_1} + \boldsymbol{\Gamma}_{\mathrm{N}_2} + \boldsymbol{\Gamma}_{\mathrm{F}} + \boldsymbol{\Gamma}_{\mathrm{W}} = \mathbf{0}.$$

In equilibrium,

$$|\mathbf{F}| \leq \mu|\mathbf{N}_1|,$$

where μ is the coefficient of static friction.

By the choice of origin, two of the position vectors are zero. The other position vectors are ◀ Solve equation(s) ▶

$$\mathbf{r}_{\mathrm{W}} = -l\cos\tfrac{\pi}{3}\,\mathbf{i} + l\sin\tfrac{\pi}{3}\,\mathbf{j} = -\tfrac{1}{2}l\,\mathbf{i} + \tfrac{\sqrt{3}}{2}l\,\mathbf{j},$$
$$\mathbf{r}_{\mathrm{N}_2} = -\sqrt{3}r\cos\tfrac{\pi}{3}\,\mathbf{i} + \sqrt{3}r\sin\tfrac{\pi}{3}\,\mathbf{j} = -\tfrac{\sqrt{3}}{2}r\,\mathbf{i} + \tfrac{3}{2}r\,\mathbf{j}.$$

The forces are given by

$$\mathbf{N}_1 = |\mathbf{N}_1|\,\mathbf{j},$$
$$\mathbf{N}_2 = |\mathbf{N}_2|\cos\tfrac{\pi}{6}\,\mathbf{i} + |\mathbf{N}_2|\sin\tfrac{\pi}{6}\,\mathbf{j} = \tfrac{\sqrt{3}}{2}|\mathbf{N}_2|\,\mathbf{i} + \tfrac{1}{2}|\mathbf{N}_2|\,\mathbf{j},$$
$$\mathbf{F} = |\mathbf{F}|\,(-\mathbf{i}),$$
$$\mathbf{W} = mg\,(-\mathbf{j}).$$

So the non-zero torques are

$$\boldsymbol{\Gamma}_{\mathrm{W}} = \left(-\tfrac{1}{2}l\,\mathbf{i} + \tfrac{\sqrt{3}}{2}l\,\mathbf{j}\right) \times (-mg\,\mathbf{j}) = \tfrac{1}{2}mgl\,\mathbf{k},$$
$$\begin{aligned}\boldsymbol{\Gamma}_{\mathrm{N}_2} &= \left(-\tfrac{\sqrt{3}}{2}r\,\mathbf{i} + \tfrac{3}{2}r\,\mathbf{j}\right) \times \left(\tfrac{\sqrt{3}}{2}|\mathbf{N}_2|\,\mathbf{i} + \tfrac{1}{2}|\mathbf{N}_2|\,\mathbf{j}\right)\\ &= -\tfrac{\sqrt{3}}{4}r\,|\mathbf{N}_2|\,\mathbf{i}\times\mathbf{j} + \tfrac{3\sqrt{3}}{4}r\,|\mathbf{N}_2|\,\mathbf{j}\times\mathbf{i}\\ &= -\sqrt{3}r\,|\mathbf{N}_2|\,\mathbf{k}.\end{aligned}$$

Substituting into the second equilibrium equation and resolving in the $\mathbf{k}$-direction gives

$$-\sqrt{3}r\,|\mathbf{N}_2| + \tfrac{1}{2}mgl = 0.$$

Therefore

$$|\mathbf{N}_2| = \frac{\frac{1}{2}mgl}{\sqrt{3}r} = \frac{mgl}{2\sqrt{3}r}.$$

Resolving the first equilibrium equation in the $\mathbf{i}$-direction gives

$$\tfrac{\sqrt{3}}{2}|\mathbf{N}_2| - |\mathbf{F}| = 0,$$

so

$$|\mathbf{F}| = \frac{\sqrt{3}}{2} \times \frac{mgl}{2\sqrt{3}r} = \frac{mgl}{4r}.$$

Resolving in the **j**-direction gives

$$|\mathbf{N}_1| + \tfrac{1}{2}|\mathbf{N}_2| - mg = 0,$$

so

$$|\mathbf{N}_1| = mg - \frac{1}{2} \times \frac{mgl}{2\sqrt{3}r} = mg\left(1 - \frac{l}{4\sqrt{3}r}\right).$$

Hence, in equilibrium, from the friction condition we have

$$|\mathbf{F}| = \frac{mgl}{4r} \leq \mu|\mathbf{N}_1| = \mu mg\left(1 - \frac{l}{4\sqrt{3}r}\right),$$

so

$$\frac{l}{4r} \leq \mu\left(1 - \frac{l}{4\sqrt{3}r}\right) = \mu\left(\frac{4\sqrt{3}r - l}{4\sqrt{3}r}\right).$$

Now, we are given that the distance OA is greater than l, and as we saw above, the distance OA is $x = \sqrt{3}r$. So we have $\sqrt{3}r > l$. Therefore we certainly have $4\sqrt{3}r - l > 0$, hence we can rearrange the equation above as

$$\mu \geq \frac{\sqrt{3}l}{4\sqrt{3}r - l}.$$

◀ Interpret solution ▶

The model predicts that the least coefficient of static friction that ensures equilibrium is $\sqrt{3}l/(4\sqrt{3}r - l)$.

Unit 3

Dynamics

Introduction

Think of an object that is moving in some way. It might be a car accelerating on a motorway, a tennis ball flying through the air, a comet hurtling through space, or a pendulum swinging to and fro. Why does the object move as it does? How will it move in the future? To what extent can you influence its motion? Questions like these are very important from a practical point of view. The control that the human race exerts over the environment depends, to a large extent, on our ability to find the right answers. For example, in Figure 1, can we predict whether the comet will collide with the Earth in time for something to be done about it?

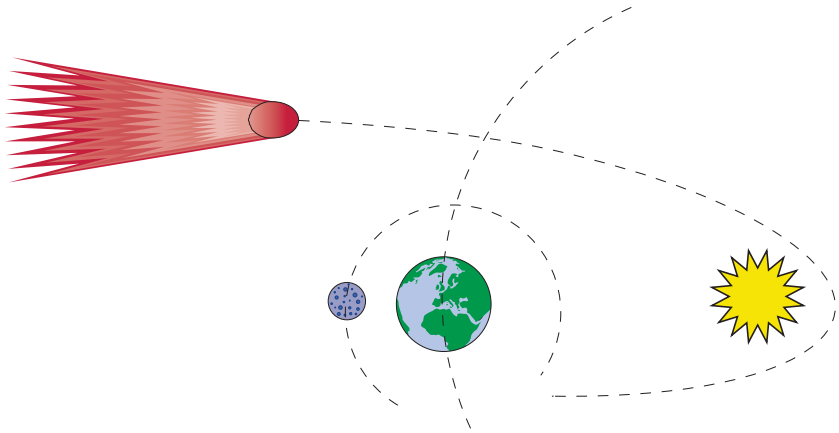


Figure 1 Collision course?

Fortunately, the systematic, organised body of knowledge called *mechanics*, which you began to study in Unit 2, provides answers to many questions about motion and the forces that cause it.

Unit 2 concentrated on whether objects will stay put; this unit looks at moving objects, that is, **dynamics**. The starting point for all of Newtonian mechanics is Newton's laws of motion, which were stated in words in the Introduction of Unit 2. To apply these laws, we need to be able to translate them into useful mathematical equations – which is one of the aims of this unit.

You should bear in mind that this unit and Unit 2 provide only an introduction to the ideas of mechanics. However, the framework of mechanics presented in these units is of great significance as it provides the foundations from which, in later units, more complex ideas are developed.

Section 1 of this unit is concerned with concepts like position, velocity and acceleration, which describe the way an object moves. Section 2 discusses Newton's laws of motion, which predict the motion of an object when the forces acting on the object are known. Section 3 shows how Newton's second law of motion can be used to predict the motion of objects. Section 4 is concerned with modelling some of the forces that occur in nature, which enables more realistic situations to be analysed.

In Section 5, we turn to motion in more than one dimension. In Unit 2 you met vectors – quantities with magnitude and direction. In modelling motion in more than one dimension, the role of vectors is more crucial, as it was in Unit 2 when modelling forces in equilibrium in two and three dimensions.

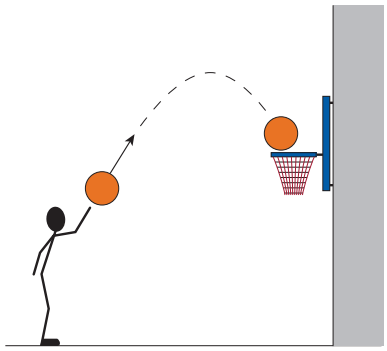


Figure 2 A basketball in flight

A thrown object, such as a basketball (see Figure 2) or a shot, from the point of release until it hits the ground, is subject only to the force of gravity and to any force exerted on it by the air (broadly referred to as *air resistance*). Such objects are examples of **projectiles**, where we know the forces acting on the object and want to deduce the path of the object through space. Of course, such thrown objects may happen to execute motion that is purely vertical (straight up and straight down), but our interest here is in cases where the motion is horizontal as well as vertical. Athletic and sporting activities provide a wide variety of examples of projectile motion. As well as the throwing or striking of objects such as basketballs or golf balls, sports may involve humans themselves acting as projectiles, as in diving, long-jumping and ski-jumping.

Here we will deal with models of projectiles without air resistance. Section 5 examines a variety of aspects of such motion.

In many of the examples considered in this unit, the size of the moving object is of significance. For example, in a football free kick, the ball is often given a spinning motion to make it swerve through the air. We will not consider such aspects of motion here, however. In this unit we consider only objects modelled as particles.

Throughout this unit we make use of Newton's second law of motion in vector form, given by

$$\mathbf{F} = m\mathbf{a},$$

Remember to underline vectors in handwritten work.

where the total force $\mathbf{F}$ and the acceleration $\mathbf{a}$ are vectors. In Sections 1 to 4, this law is used almost exclusively in one dimension. In Section 5, we will be concerned with its use in more than one dimension. If the position vector of a particle is known as a function of time, we can obtain its velocity and acceleration vectors by differentiation. In more than one dimension, we use the position vector $\mathbf{r}(t)$ to represent the position of a particle at time t. To differentiate $\mathbf{r}(t)$ with respect to t, we differentiate separately each of the functions giving its $\mathbf{i}$-, $\mathbf{j}$- and $\mathbf{k}$-components. We usually employ the notation $\dot{\mathbf{r}}(t)$ (or just $\dot{\mathbf{r}}$) for the velocity $\mathbf{v} = d(\mathbf{r}(t))/dt$. Similarly, we usually write $\ddot{\mathbf{r}}(t)$ (or just $\ddot{\mathbf{r}}$) for the second derivative of $\mathbf{r}(t)$ that gives the acceleration $\mathbf{a}$.

1 Describing motion

This section is devoted to describing the motion of objects modelled as particles. In Subsection 1.1 the motion is described by giving a position vector at each instant in time. The ideas in Subsection 1.1 apply whether the particle moves along a straight line or along a curve of some sort; in Subsection 1.2, however, and for Sections 2 to 4, only motion along a straight line is considered.

The subject matter of this section, the description of motion, is often referred to by the technical term *kinematics*.

1.1 Motion of a particle

The motion of a real object, say a leaf that is falling to the ground, is very difficult to describe exactly. The leaf may rotate, bend or vibrate while moving along a complicated path in three-dimensional space. And its motion may be affected by the presence of other moving objects, such as other falling leaves. It would be foolhardy to try to meet all of these difficulties head-on, so we start by making a number of simplifications, some of which will be relaxed later.

The motion of rigid bodies is discussed in Unit 21, and many-particle systems are discussed in Unit 19.

Simplification 1 Objects will be modelled as particles.

Simplification 2 Only the motion of *single* particles is considered.

Modelling objects as particles means that we neglect an object's size and internal structure. Neglecting an object's size means that its location at any given time may be described by a single point in space, and that its motion may be described by a single curve. Neglecting the internal structure of an object, and hence also any internal motion, amounts to saying that the curve described in time by the particle gives the only information of interest about the way in which the object moves. In mathematical terms, this means that a particle's motion is completely described by its *position vector* $\mathbf{r}$, relative to some fixed origin O, at those times t in the time interval of interest.

The representation of the motion of a particle by a position vector that changes with time leads naturally to the representation of such motion by a *vector function*.

A **vector function** $\mathbf{r}(t)$ of some variable t is a vector whose components are functions of t.

Unlike the components of a vector, which are constant, the components of a vector function vary as the independent variable changes. In the case of the motion of a particle, a vector function $\mathbf{r}(t)$, whose components at each time t represent the position vector $\mathbf{r}$ of the particle at that time, completely describes the particle's motion. The idea is illustrated in Figure 3 and in the following example of two-dimensional motion.

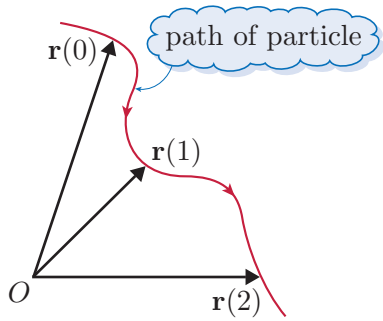


Figure 3 The path of a particle described by the vector function $\mathbf{r}(t)$

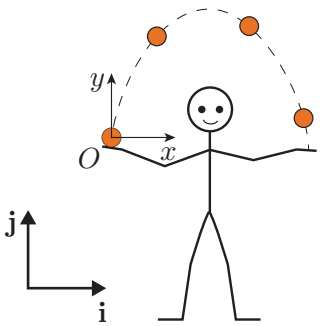


Figure 4 A juggler with a ball

Example 1

A juggler throws a ball from one hand to the other, in a vertical plane, as shown in Figure 4. The ball is modelled as a particle, and its motion, with respect to the horizontal and vertical axes shown in Figure 4, is described by the two-dimensional vector function

$$\mathbf{r}(t) = 1.5t\mathbf{i} + t(4 - 5t)\mathbf{j} \quad (0 \leq t \leq 1),$$

where distances are measured in metres, and time t is measured in seconds after the ball was thrown. The origin O is the juggler's right hand just as he throws the ball, which occurs at time $t = 0$.

Let $x(t)$ be the component of $\mathbf{r}(t)$ in the $\mathbf{i}$-direction, that is, $x(t) = \mathbf{r}(t) \cdot \mathbf{i} = 1.5t$. Similarly, let $y(t)$ be the component of $\mathbf{r}(t)$ in the $\mathbf{j}$-direction, that is, $y(t) = \mathbf{r}(t) \cdot \mathbf{j} = t(4 - 5t)$. These give the horizontal distance travelled and the height of the ball, respectively, at time t.

(a) Calculate $x(t)$ and $y(t)$ at times $t = 0, 0.2, 0.4, 0.6, 0.8, 1$, and sketch the graphs of $y(t)$ against t and against $x(t)$. Comment on what the graphs represent.

(b) Using your graphs, or otherwise, answer the following questions.

(i) How high does the ball go, and what is its position at its highest point?

(ii) Does the juggler catch the ball?

Solution

(a) The values are tabulated in Table 1, and the graphs are shown in Figure 5.

Table 1

t	0	0.2	0.4	0.6	0.8	1
$x(t)$	0	0.3	0.6	0.9	1.2	1.5
$y(t)$	0	0.6	0.8	0.6	0	−1

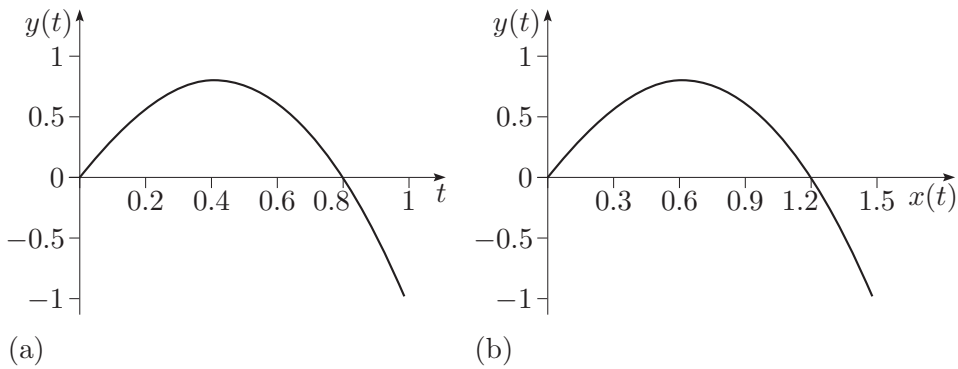


Figure 5 Graphs of (a) height against time, (b) height against horizontal distance

Figure 5(a) is a distance–time graph, representing the *height* $y(t)$ of the ball as the time t varies over the interval $[0, 1]$. Figure 5(b) represents the *position* of the ball in the (x, y)-plane as t varies; the coordinates $(x(t), y(t))$ of points on this curve are the components of the vector function $\mathbf{r}(t)$ describing the motion of the ball. In this case, since $x(t) = 1.5t$ is just a multiple of t, the two curves appear identical – the only difference is in the scales on the horizontal axes. However, the curves *represent* different things.

You may recognise $(x(t), y(t))$ as a *parametrisation* of the path of the ball, in which the Cartesian coordinates of the ball are expressed in terms of another variable (in this case time t).

(b) (i) The quadratic function $y(t) = t(4 - 5t)$ represents a parabola with maximum value in the interval $[0, 0.8]$ (see Figure 5(a)). From the symmetry of the parabola, this maximum must occur at $t = 0.4$, at which time $y(t) = 0.8$ (see Table 1). So the ball reaches a maximum height of 0.8 m above the juggler's hand after 0.4 s. At this maximum height, $x(t) = 0.6$ (see Table 1).

So the ball's position at its maximum height is given by the coordinates $(0.6, 0.8)$ (see Figure 5(b)), or equivalently, by the position vector $\mathbf{r}(0.4) = 0.6\mathbf{i} + 0.8\mathbf{j}$.

(ii) The juggler does not catch the ball because the data indicate that the ball continues to travel downwards until it is one metre below the juggler's left hand (see Table 1 and Figure 5).

In general, motion in three-dimensional space is represented by a three-dimensional vector function. However, because the motion in Example 1 was in a plane, it was possible by careful choice of axes to represent that motion in three-dimensional space by a two-dimensional vector function. A similarly careful choice of axes can enable certain types of motion in three-dimensional space to be represented by a *one*-dimensional vector function, as the following exercise illustrates.

Exercise 1

An ice hockey player aims to hit the puck towards the goal, as shown in Figure 6, but misses. The puck hits the back wall and bounces straight back, and is then hit by a second player who (incredibly) hits it along the same path as the first player. Let the origin O be the point of impact of the puck by the first player, which occurs at time $t = 0$. The motion of the puck, with respect to the axis shown in Figure 6, is described by the one-dimensional vector function $\mathbf{r}(t) = x(t)\,\mathbf{i}$, where the function $x(t)$ is defined by

$$x(t) = \begin{cases} 5t, & 0 \le t < 0.4, \\ 4 - 5t, & 0.4 \le t < 1, \\ 6t - 7, & 1 \le t \le 1.5, \end{cases}$$

where x is measured in metres, and t is measured in seconds.

(a) Sketch the graph of the function $x(t)$.

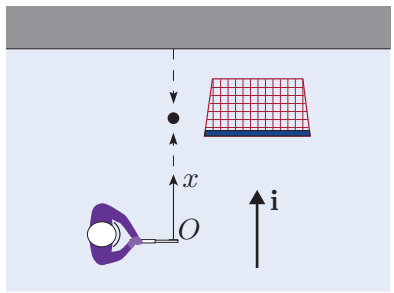


Figure 6 The path of the puck

(b) Using your graph, or otherwise, answer the following questions.

(i) How far is the first player from the back wall when the puck is struck?

(ii) Does the second player hit the puck from a position closer to or further from the back wall than the first player?

(iii) Does the second player give the puck more speed than the first player?

You saw in Unit 2 that some large static objects can be modelled as particles.

In Example 1 and Exercise 1, objects were modelled as particles. The question arises as to when this is appropriate. It seems obvious that small objects can be modelled as particles, but what about large objects such as the Earth? The answer is that it depends on the context. For example, to calculate the Earth's orbit around the Sun, it is permissible to model the Earth as a particle. One of Newton's great achievements was to realise that it is appropriate to model the Earth as a particle in this context.

The Earth's diameter (about 13 000 km) is small compared with the Earth–Sun distance (about 1.5×10^8 km).

Whether a particle model will be satisfactory is not just a question of size. For example, if a ball is placed on a rough sloping table, then it will roll down the slope. A particle model could be used to describe the path of the ball's centre, but it would not be adequate to keep track of the rolling motion that takes place about the centre. Note that this inadequacy of the particle model occurs regardless of the ball's size, since the same consideration would apply to a football, a tennis ball, a marble or a ball-bearing. However, if the ball is not rotating, or if it is assumed that the rotations can be ignored, then the particle model is appropriate.

You might think from this last example that the particle model is of very limited use for moving objects, but in fact the example hints at how this model can be extended. You will see in Unit 19 that the motion of an object can be described well by specifying:

You met the idea of centre of mass in Unit 2. For the symmetric objects considered in that unit, the centre of mass is the geometric centre.

- the motion of the *centre of mass* of an object
- the motion of the whole object relative to its centre of mass.

As you will see in Unit 19, the motion of the centre of mass may be predicted by considering a particle of the same mass as the object, placed at the point defined by the centre of mass, and subjected to all of the forces that act on the object. So even in this more complicated situation, the concept of a particle is important. Alternatively, it may be appropriate to think of an object as being composed of a number of elements, each of which can be modelled individually as a particle.

Position, velocity and acceleration

As in Unit 1, we may sometimes wish to consider position as a variable rather than as a function, and write $\mathbf{r}$ rather than $\mathbf{r}(t)$.

We have seen how a particle's motion can be completely described by a vector function $\mathbf{r}(t)$, the position vector of the particle at any given time t. For this reason, we often refer to particles as having **position** $\mathbf{r}(t)$.

However, other quantities, such as the velocity and acceleration of a particle, are often of more interest than its position. For example, an

aggressive motorist might be proud of his acceleration away from traffic lights, whereas a police officer would probably be more interested in the motorist's velocity. These quantities can be calculated directly from a particle's position $\mathbf{r}(t)$ by differentiating. But before we can do that, we need to define the derivative of a vector function.

The **derivative** of a vector function $\mathbf{r}(t)$ whose components are smooth functions of t, is the vector function

$$\frac{d\mathbf{r}(t)}{dt} = \lim_{h\to 0}\left(\frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}\right). \tag{1}$$

The derivative $d\mathbf{r}(t)/dt$ is often written as $\mathbf{r}'(t)$ or, where t represents time, as $\dot{\mathbf{r}}(t)$. Sometimes it is written more succinctly as $d\mathbf{r}/dt$, $\mathbf{r}'$ or $\dot{\mathbf{r}}$.

Compare this with the definition of the derivative of a function as

$$\frac{dy}{dx} = \lim_{h\to 0}\left(\frac{y(x+h)-y(x)}{h}\right).$$

Newton's notation $\dot{\mathbf{r}}(t)$ is commonly used in mechanics.

This definition makes use of the concept of the **limit** of a vector function. As you might expect, the limit of a vector function $\mathbf{f}(h)$ as $h \to 0$ is the vector function whose components are the limits, as $h \to 0$, of the components of $\mathbf{f}(h)$.

Now, as you will recall, velocity is defined to be rate of change of position, so the definition of the derivative of a vector function can be used to define the **velocity** of a particle with position $\mathbf{r}(t)$ as

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt}.$$

We sometimes consider velocity as a variable rather than as a function, and write $\mathbf{v}$ rather than $\mathbf{v}(t)$.

The **speed** of the particle, which as you will recall is defined to be the magnitude of the velocity, is therefore given by $|\mathbf{v}(t)|$.

To find the derivative of a vector function, we make use of the following theorem.

Theorem 1

If a vector function $\mathbf{r}(t)$ has the component form

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k},$$

where $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are the (fixed-direction) Cartesian unit vectors, then its derivative is given by

$$\frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}. \tag{2}$$

Exercise 2

For each of the particles whose positions are given, calculate the velocity and speed of the particle at $t = 1$, correct to two decimal places.

(a) $\mathbf{r}(t) = t^2\mathbf{i} + 10t\mathbf{j}$ (b) $\mathbf{r}(t) = (\sin t)\mathbf{i} + (\cos t)\mathbf{j} + t\mathbf{k}$

In many situations, velocity is less important than *changes* in velocity. For example, if you are on board a train travelling at a steady speed, you may not even notice that you are moving. You will have no difficulty in, say, drinking a cup of tea. However, this operation becomes more hazardous if the driver changes the velocity of the train by putting on the brakes! Similarly, if the train goes round a bend at constant speed, you will notice the change in velocity (hot tea in your lap again!). In both cases, the rate of change of velocity is an important factor.

Now, you will recall that just as velocity is defined to be rate of change of position, so acceleration is defined to be rate of change of velocity. Therefore, along similar lines to the definition of the velocity of a particle, the **acceleration** of a particle with velocity $\mathbf{v}(t)$ is defined as

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt}.$$

As with position and velocity, we sometimes write $\mathbf{a}$ rather than $\mathbf{a}(t)$.

Notice that the definition of acceleration as the rate of change of *velocity* differs from its common everyday meaning as the rate of increase of *speed*. For example, consider the train moving at constant speed around a bend mentioned above. Its speed is constant, so it is not accelerating in the everyday sense. However, it is accelerating in a mathematical sense, because its velocity is changing direction. Similarly, in the everyday sense, the braking train is not accelerating but decelerating, in that its speed is decreasing. However, in a mathematical sense it is accelerating since its velocity is changing.

It is important to understand the difference between the mathematical and everyday meanings of acceleration, and also to be able to interpret the meaning of the components of the vectors defining the position, velocity and acceleration of a particle. The following exercise should help you to do this.

Exercise 3

Three particles, A, B and C, are moving along three different straight lines. In each case, the straight line is chosen as the x-axis, so the vectors defining the position, velocity and acceleration of the three particles have only x-components. The graphs of the x-components $x(t)$ of the position of particle A, $\dot{x}(t)$ of the velocity of particle B, and $\ddot{x}(t)$ of the acceleration of particle C are shown in Figure 7. Use the graphs to answer the questions that follow.

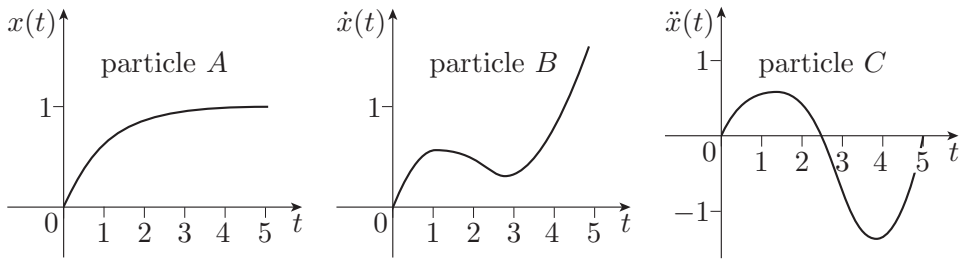


Figure 7 Graphs against t of x for particle A, $\dot{x}$ for particle B, and $\ddot{x}$ for particle C

(a) When is particle A travelling fastest?

(b) Does particle B change its direction of motion in the time interval shown?

(c) Particle C starts from rest at $x = 0$ at time $t = 0$. Is particle C momentarily stationary at any later time in the time interval shown?

(*Hint*: Think about the meaning of the area beneath a curve.)

1.2 One-dimensional motion

The ideas in Subsection 1.1 apply to motion in three-dimensional space generally. For the rest of this section and in the next three sections, we make the simplification that the motion will be in one dimension, so that the fundamental theory is not obscured.

Restriction to one-dimensional motion still allows a wide range of situations to be covered. One such situation, which is analysed fully in Section 3, is the following.

> If a marble is dropped from the Clifton Suspension Bridge, how long does it take to fall into the River Avon below? What is its velocity just before it hits the water?

You met this situation in Unit 1.

To answer these questions, we need expressions for the position and velocity of the falling marble. It is reasonable to assume that the path of the marble is a straight line; so, under this assumption, although the marble is moving in three-dimensional space, this is a one-dimensional problem.

With badly chosen axes, all three components of the position vector of the marble could be changing. (Imagine fixing your axes on a car as it crossed the bridge at constant speed, or simply axes inclined to the vertical.) What makes this a one-dimensional problem is that if we choose the axes well, then only one component of the position vector is changing. Normally this axis is labelled as the x-axis, with the result that the vector has only an $\mathbf{i}$-component. Furthermore, if the situation being considered has an obvious direction of positive motion, then it is usual to choose this direction as the direction of the positive x-axis. So in the case of the Clifton Suspension Bridge problem (see Figure 8), it is sensible to choose the x-axis Ox pointing vertically downwards, so that the marble moves in the positive x-direction.

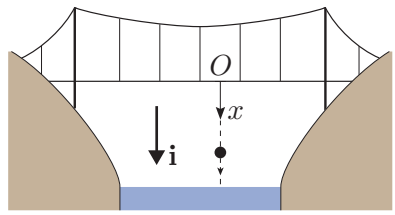


Figure 8 A marble dropped from the Clifton Suspension Bridge

For one-dimensional problems in which the x-axis corresponds to the direction of motion, the position, velocity and acceleration of the particle can be expressed as

$$\mathbf{r}(t) = x(t)\,\mathbf{i}, \quad \mathbf{v}(t) = \frac{dx(t)}{dt}\,\mathbf{i}, \quad \mathbf{a}(t) = \frac{d^2x(t)}{dt^2}\,\mathbf{i},$$

respectively. So for such problems, the position of the particle can be described by the function $x(t)$, the velocity by the function $v(t) = dx(t)/dt$, and the acceleration by the function $a(t) = dv(t)/dt = d^2x(t)/dt^2$.

As with $\mathbf{r}$, $\mathbf{v}$ and $\mathbf{a}$, we will sometimes find it convenient to write x, v and a instead of $x(t)$, $v(t)$ and $a(t)$.

In fact, for one-dimensional motion *along a given x-axis*, the vector quantities position, velocity and acceleration are completely described by the corresponding (scalar) functions $x(t)$, $v(t)$ and $a(t)$, respectively. For example, the magnitude of the velocity is given by $|v(t)|$, its orientation by the given x-axis, and its sense by the sign of $v(t)$. (Remember from Unit 2 that the direction of a vector comprises its orientation and its sense.) Therefore for the one-dimensional motion of an object along a given x-axis, we can – and frequently will – refer to $x(t)$ (or x), $v(t)$ (or v) and $a(t)$ (or a) as the position, velocity and acceleration of the object.

In most of the examples and exercises in this unit so far, the position of the object was known. It is more usual for the forces acting on the object to be known. As you will see in Section 2, this gives information about the acceleration of the object. So in the following examples and exercises we practise finding the position and velocity of a particle, given its acceleration.

Example 2

A particle is moving in a straight line along the x-axis. The acceleration $\mathbf{a}$ of the particle at time t is given by

$$\mathbf{a}(t) = (12t^2 + 2)\mathbf{i}.$$

Note that the 'initial' condition for a problem does not have to be at $t = 0$.

After 1 second, the particle is 3 metres from the origin and has velocity $2\,\mathrm{m\,s^{-1}}$.

(a) Find the velocity $\mathbf{v}(t)$ and the position $\mathbf{r}(t)$ of the particle.

(b) Find the velocity and position of the particle at time $t = 2$.

Solution

(a) The acceleration is known and is a one-dimensional vector function. Using the notation above, we have

$$a(t) = \mathbf{a}(t) \cdot \mathbf{i} = 12t^2 + 2.$$

In general, when trying to find the position of a particle given its acceleration, it is often easier to first find its velocity and then find its position, and thus solve two first-order differential equations, rather than to find its position directly by solving a second-order differential equation.

Since $a(t) = dv(t)/dt$, we have the first-order differential equation

$$\frac{dv}{dt} = 12t^2 + 2.$$

This equation can be solved by direct integration, giving

$$v = \int (12t^2 + 2)\,dt = 4t^3 + 2t + A,$$

where A is a constant.

To find the value of the constant A, we use the fact that $v = 2$ when $t = 1$, so $2 = 4 + 2 + A$, giving $A = -4$. This gives

$$v(t) = 4t^3 + 2t - 4.$$

Since $v(t) = dx(t)/dt$, we have the first-order differential equation

$$\frac{dx}{dt} = 4t^3 + 2t - 4.$$

Solving this equation by direct integration gives

$$x = \int (4t^3 + 2t - 4)\, dt = t^4 + t^2 - 4t + B,$$

where B is a constant.

The initial condition $x = 3$ when $t = 1$ can be used to find the constant B, so $3 = 1 + 1 - 4 + B$, giving $B = 5$. This gives

$$x(t) = t^4 + t^2 - 4t + 5.$$

So the velocity and position of the particle are

$$\mathbf{v}(t) = (4t^3 + 2t - 4)\mathbf{i}, \quad \mathbf{r}(t) = (t^4 + t^2 - 4t + 5)\mathbf{i}.$$

(b) When $t = 2$, the velocity and position have values

$$\mathbf{v}(2) = (4 \times 2^3 + 2 \times 2 - 4)\mathbf{i} = 32\mathbf{i},$$
$$\mathbf{r}(2) = (2^4 + 2^2 - 4 \times 2 + 5)\mathbf{i} = 17\mathbf{i}.$$

Hence at time 2 seconds, the particle has position 17 metres along the positive x-axis and velocity 32 metres per second in the direction of the positive x-axis.

Exercise 4

A particle is moving in a straight line along the x-axis. At time t the particle has an acceleration given by

$$\mathbf{a}(t) = (18t - 20)\mathbf{i} \quad (t \geq 0).$$

Initially, at $t = 0$, the particle has position $\mathbf{r}(0) = 7\mathbf{i}$ and velocity $\mathbf{v}(0) = 3\mathbf{i}$.

Find the position, velocity and speed of the particle at time $t = 10$.

Exercise 5

A particle is moving in a straight line along the x-axis. At time t the particle has an acceleration given by

$$\mathbf{a}(t) = ge^{-kt}\mathbf{i} \quad (t \geq 0),$$

where g and k are positive constants. Initially, at $t = 0$, the particle is at the origin ($\mathbf{r}(0) = \mathbf{0}$) and is at rest ($\mathbf{v}(0) = \mathbf{0}$).

Find the velocity and position of the particle as vector functions.

In general, an **equation of motion** is any equation relating two or more of acceleration, velocity, position and time. The rest of this subsection is devoted to problems involving the solution of equations of motion. We begin with the case of constant acceleration, which occurs frequently.

Example 3

A particle moves in a straight line along the x-axis with constant acceleration $\mathbf{a}(t) = a_0\mathbf{i}$ ($a_0 \neq 0$). The particle starts from $x = x_0$ at time $t = 0$ with initial velocity $\mathbf{v}(0) = v_0\mathbf{i}$.

The subscript 0 is used to distinguish the constants a_0 and v_0 from the variables a and v.

(a) Show that the velocity vector and position vector of the particle are given by $\mathbf{v}(t) = v(t)\,\mathbf{i}$ and $\mathbf{r}(t) = x(t)\,\mathbf{i}$, where

$$v(t) = v = v_0 + a_0t,$$
$$x(t) = x = x_0 + v_0t + \tfrac{1}{2}a_0t^2.$$

(b) By eliminating t between these two equations, show that

$$v^2 = v_0^2 + 2a_0(x - x_0).$$

Solution

(a) We have

$$a(t) = \frac{dv}{dt} = a_0,$$

which on integration yields

$$v = \int a_0\,dt = a_0t + A,$$

where A is a constant.

The initial condition $v(0) = v_0$ gives $A = v_0$, so the velocity is given by

$$v(t) = v = v_0 + a_0t. \tag{3}$$

Hence

$$\frac{dx}{dt} = v_0 + a_0t,$$

from which

$$x = \int (v_0 + a_0t)\,dt = v_0t + \tfrac{1}{2}a_0t^2 + B,$$

where B is a constant.

Since the particle starts at $x = x_0$, we have $B = x_0$, so the position is given by

$$x(t) = x = x_0 + v_0t + \tfrac{1}{2}a_0t^2. \tag{4}$$

(b) Rearranging equation (3) gives, since $a_0 \neq 0$,

$$t = \frac{v - v_0}{a_0}.$$

Substituting this into equation (4) yields

$$x = x_0 + v_0\left(\frac{v - v_0}{a_0}\right) + \tfrac{1}{2}a_0\left(\frac{v - v_0}{a_0}\right)^2.$$

Multiplying through by $2a_0$ and expanding the brackets gives

$$2a_0(x - x_0) = 2(v_0v - v_0^2) + (v^2 - 2v_0v + v_0^2) = v^2 - v_0^2,$$

which can be rearranged to give

$$v^2 = v_0^2 + 2a_0(x - x_0),$$

as required.

The results of Example 3 can be summarised as follows.

Constant acceleration

If a particle is moving in a straight line along the x-axis with constant acceleration $\mathbf{a}(t) = a_0\mathbf{i}$, and at time $t = 0$ it has initial position $\mathbf{r}(0) = x_0\mathbf{i}$ and initial velocity $\mathbf{v}(0) = v_0\mathbf{i}$, then the components of acceleration, velocity and position along the x-axis are given, respectively, by

$$a = a_0, \tag{5}$$

$$v = v_0 + a_0t, \tag{6}$$

$$x = x_0 + v_0t + \tfrac{1}{2}a_0t^2. \tag{7}$$

Furthermore,

$$v^2 = v_0^2 + 2a_0(x - x_0). \tag{8}$$

We know that, by definition, $a = dv/dt = d^2x/dt^2$. There is a useful alternative expression for a that can be derived using the chain rule:

$$a = \frac{dv}{dt} = \frac{dv}{dx}\frac{dx}{dt} = \frac{dv}{dx}v.$$

Alternative expressions for $a(t)$

$$a = \frac{dv}{dt} = \frac{d^2x}{dt^2} = v\frac{dv}{dx}. \tag{9}$$

The formula $a = v\,dv/dx$ can be used to obtain equation (8) directly, without having to find v and x first, as the following exercise asks you to demonstrate.

Exercise 6

A particle moves in a straight line along the x-axis with constant acceleration $\mathbf{a}(t) = a_0\mathbf{i}$. Initially, at time $t = 0$, the particle is at $x = x_0$ and has velocity $\mathbf{v}(0) = v_0\mathbf{i}$.

Use the relationship $a = v\,dv/dx$ to show that

$$v^2 = v_0^2 + 2a_0(x - x_0).$$

In general, given an equation of motion relating acceleration to one or more of velocity, position and time, we want to obtain an equation relating velocity to position and/or time, or an equation relating position to time. In the case of one-dimensional motion, we can do this by using one of

$$a = \frac{dv}{dt}, \quad a = \frac{d^2x}{dt^2}, \quad a = v\,\frac{dv}{dx}$$

to substitute for a and then solving the resulting differential equation. The following exercise asks you to decide which formula for a provides the most appropriate substitution in a variety of typical cases.

Exercise 7

How would you use the above formulas to substitute for a in the following equations of motion in order to obtain the specified information? (You are not expected to solve the resulting equations.)

(a) The x-component of the equation of motion is $a = \cos t$; it is required to find velocity and position in terms of time.

(b) The x-component of the equation of motion is $a = -x$; it is required to find a relationship between velocity and position.

(c) The x-component of the equation of motion is $a = -2x - 3v + \cos t$; it is required to find position in terms of time.

The models developed in this section are used in practice as shown in the next exercise.

Exercise 8

The data in Table 2, taken from the United Kingdom *Highway Code*, show the shortest stopping distances of cars travelling along a straight road.

The thinking distance is defined to be the distance travelled by a car in the maximum time that it takes for an alert driver to react to a hazardous situation.

Table 2

Speed (mph)	Thinking distance (metres)	Braking distance (metres)	Overall stopping distance (metres)
20	6	6	12
30	9	14	23
40	12	24	36
50	15	38	53
60	18	55	73
70	21	75	96

The British Imperial unit for speed is miles per hour (mph), and $1\,\text{mph} \simeq 0.447\,\text{m}\,\text{s}^{-1}$.

(a) The data in Table 2 are not from an experiment; they are the predictions of models. Your task is to discover what models were used.

(i) What model (using SI units) was used to obtain the thinking distance data?

(*Hint*: Think about the speed of the car before and after the thinking phase, and use the constant acceleration formula $x = x_0 + v_0 t + \frac{1}{2}a_0 t^2$.)

(ii) What model (using SI units) was used to obtain the braking distance data?

(*Hint*: Think about the speed of the car before and after the braking phase, and use the constant acceleration formula $v^2 = v_0^2 + 2a_0 x$.)

(b) Use your models from part (a) to predict the overall stopping distance (in metres) for a speed of 45 mph.

2 A theory of motion

Section 1 introduced the basic concepts of position, velocity and acceleration that are needed to *describe* motion. In this section, two concepts introduced in Unit 2, *force* and *mass*, enable us to go beyond the mere description of motion and formulate laws *predicting* what motions take place.

At first sight it might seem that a different set of rules of motion would be required for each type of object – one set for tennis balls, another set for planets, and so on. Fortunately, there is a simple underlying pattern. Newton was able to see beyond individual cases, and his three laws of motion form a framework, or theory, for predicting the motion of all objects.

Newton's three laws of motion were stated in the Introduction of Unit 2.

You met the idea of friction in Unit 2. It is discussed further in Subsection 4.1 of this unit.

Our instinctive ideas about motion are shaped by the presence of friction in almost all things in our everyday lives. Aristotle (384–322 BC) constructed a theory out of this experience that turned out to be completely wrong. This theory and the subsequent development of ideas of motion led eventually to Newton's laws. Consider the following thought experiment to imagine motion without friction.

Consider a toboggan on a horizontal icy surface such as a frozen lake. Left undisturbed, the toboggan remains static; it must be pushed or pulled in some way if it is to be set in motion, that is, a force must act on the toboggan. However, if you give the toboggan a push and then release it, the toboggan will move across the ice at almost constant speed in the direction in which it has been pushed.

Once the toboggan has been released, the force of the push ceases to act on it.

This suggests that under ideal (i.e. frictionless) conditions, the following applies:

> in the absence of a force, the toboggan remains at rest or moves with constant speed in a straight line.

Air resistance is discussed in Subsection 4.2.

In real life, the toboggan does eventually slow down, partly due to air resistance and partly due to friction between the toboggan runners and the ice. In competitive tobogganing, the tobogganers go to great lengths to reduce these resistive forces (i.e. forces resisting motion) by streamlining the toboggan to reduce air resistance, and waxing the runners to reduce friction.

Exercise 9

A car on a flat, straight road requires a motive force (supplied by its engine) in order to maintain a constant speed of 70 miles per hour; if the engine is switched off, then the car slows down. It might be thought from the above example that if an object is moving with constant velocity, then there is no force acting on it. Try to explain this apparent contradiction.

Returning to the example above, suppose that you apply a force by pushing the toboggan continuously. You cannot quantify this force, but the sensations in your muscles and nerves will reveal whether you are pushing gently or firmly. From experience, you know that:

> the harder you continue to push, the further and faster the toboggan moves in a given time.

This suggests that there is a link between the force that is applied and the way in which the toboggan moves.

Exercise 10

A toboggan on an icy slope may accelerate even when it is not being pushed. Try to identify the force that causes this acceleration.

Next, imagine pushing two identical toboggans, one of which is empty while the other carries a heavy person. If you apply the same force to the two toboggans, then the laden toboggan will move more sluggishly. To achieve the same motion in each case, it is necessary to apply a greater force to the laden toboggan. In other words:

> if you apply the same force to the two toboggans, the laden toboggan does not travel as far or as fast in a given time as the empty toboggan;
>
> in order for the two toboggans to move in the same way, a greater force must be applied to the laden toboggan than to the empty toboggan.

In general, it seems that three concepts are linked together:

- the *force* that is applied to an object
- the *mass* of the object
- the *motion* of the object.

Newton proposed in his book *Principia* that this link takes the form

$$\text{force} = \text{mass} \times \text{acceleration}. \tag{10}$$

The validity of this equation is shown by its success at predicting motion – nearly the whole of nineteenth-century science and engineering rested on it! It also provides a more formal definition of **force** than was given in Unit 2.

Equation (10) can be written in the vector form

$$\mathbf{F} = m\mathbf{a},$$

where $\mathbf{F}$ is the vector quantity force, m is the scalar quantity mass, and $\mathbf{a}$ is the vector quantity acceleration. This equation is the bedrock of Newtonian mechanics, and it is usually referred to as **Newton's second law**.

Strictly speaking, the argument concerning the toboggan justifies only the statement that force is *proportional* to mass times acceleration, that is, $\mathbf{F} = km\mathbf{a}$ with k some constant of proportionality. However, in the SI system of units, k is *chosen* to be 1 by an appropriate definition of the unit of force, the *newton*, which makes use of Newton's second law.

You met the newton, though not a formal definition of it, in Unit 2.

A force of magnitude one **newton** (1 N) is the force required to accelerate a mass of one kilogram at one metre per second per second (i.e. $1\,\text{N} = 1\,\text{kg}\,\text{m}\,\text{s}^{-2}$). The direction of the force is the direction in which the mass accelerates.

This is a good point at which to summarise the discussion so far in this section into precise laws comprising the foundations of Newtonian mechanics.

You met one of these laws, the law of addition of vectors (forces), in Unit 2.

Fundamental laws of Newtonian mechanics

These laws concern a **particle**, which is a mathematical model for any material object whose size and internal structure may be neglected.

The **mass** of the particle is expressed by a single positive number m. This number is an inherent property of the particle and does not depend on time, position, force or any other variable.

Law of addition of mass

If an object modelled as a particle is composed of a number of parts, then the mass m of the particle is the sum of the masses of the parts.

Law of addition of forces

If several forces act simultaneously on a particle, then the resultant force is the vector sum of the individual forces.

It is very important to take into account the directions of the individual forces as well as their magnitudes.

Newton's second law

If a particle has a constant mass m and experiences a total force $\mathbf{F}$, then its acceleration $\mathbf{a}$ is given by

$$\mathbf{F} = m\mathbf{a}. \tag{11}$$

This statement of Newton's second law applies to particles of constant mass, which are the only particles studied in this module.

Newton's first law

When $\mathbf{F}$ is zero, $\mathbf{a}$ is zero: in the absence of a force, a particle either stays permanently at rest or moves at constant velocity, that is, at a constant speed in a straight line.

Newton's first law is a special case of Newton's second law. It forms the basis of the equilibrium conditions in Unit 2.

Strictly speaking, Newton's laws hold in an *inertial frame of reference* – a frame of reference (coordinate system) that is not accelerating. It is assumed throughout this module that reference frames are inertial.

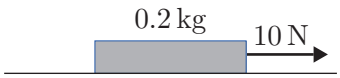


Figure 9 An object on a smooth horizontal surface

Exercise 11

If a mass of 200 grams on a smooth horizontal surface (see Figure 9) is subjected to a horizontal force of magnitude 10 newtons, what is the magnitude of the acceleration produced?

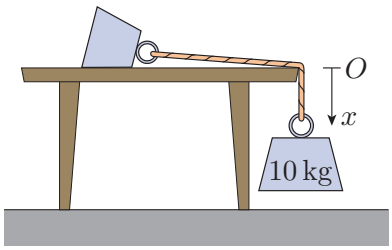


Figure 10 Two objects connected by a string

Exercise 12

An object of mass 10 kilograms is attached to a string hanging over the edge of a table, as shown in Figure 10. The other end of the string is attached to another object, on top of the table. The object hanging over the edge of the table is observed to be accelerating at $1\,\mathrm{m\,s^{-2}}$ downwards.

(a) What is the resultant force on the hanging object?

(b) Apply Newton's second law to the hanging object, and hence calculate the tension force due to the string acting on the hanging object.

Exercise 13

A fighter pilot can experience an acceleration of magnitude approximately six times the magnitude of the acceleration due to gravity before being rendered unconscious. If a fighter of mass 4000 kilograms is subjected to a force $50\,000\,\mathbf{i} + 60\,000\,\mathbf{j} + 100\,000\,\mathbf{k}$ (with magnitude in newtons) during an aerobatic manoeuvre, will the pilot remain conscious?

3 Predicting motion

Newton's second law of motion can be used to help to solve a huge variety of mechanics problems. The example below considers the motion of an object falling under gravity alone.

The steps involved in the solution are similar to those in Procedure 2 of Unit 2. As in Unit 2, the steps are highlighted by labels in the margin.

Example 4

A marble, initially at rest, is dropped from the Clifton Suspension Bridge and falls into the River Avon, 77 metres below. Assume that the bridge is fixed and the only force acting on the marble is its weight due to gravity.

In Section 4 this example will be remodelled to include air resistance.

(a) Find the time taken before the marble hits the water.

(b) Find the speed of the marble just before it hits the water.

Solution

◀ Draw picture ▶

(a) The first step is always to draw a diagram that includes all the relevant information given in the problem, as in Figure 11.

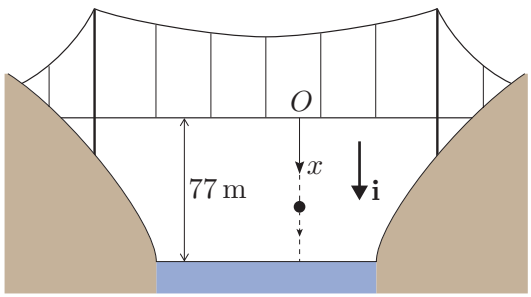


Figure 11 Sketch of the Clifton Suspension Bridge with the data marked on it

◀ Choose axes ▶

Choose an x-axis pointing vertically downwards, with its origin O at the fixed point where the object is released, as shown in Figure 11. (This makes the algebra simpler as all quantities are positive.)

◀ State assumptions ▶

The marble is modelled as a particle with its weight the only force acting on it. This is represented by the force diagram in Figure 12, in which m denotes the mass of the marble and $\mathbf{W}$ denotes its weight.

◀ Draw force diagram ▶

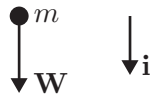


Figure 12 Force diagram for the marble

In Unit 2 the downwards vertical was usually in the $-\mathbf{j}$ direction.

The resultant force acting on the marble is $\mathbf{W} = mg\mathbf{i}$ (where, as usual, $\mathbf{i}$ is a unit vector pointing in the direction of the positive x-axis, in this case vertically downwards).

◀ Apply Newton's 2nd law ▶

The acceleration $\mathbf{a}$ is obtained by applying Newton's second law:

$$m\mathbf{a} = mg\mathbf{i}.$$

From this equation we see that only the $\mathbf{i}$-component of $\mathbf{a}$ could be varying, so we have $\mathbf{a} = a(t)\,\mathbf{i}$. So the equation becomes $ma = mg$, and dividing through by m gives

$$a = g.$$

(Notice that since m cancels out, the results apply to a marble of any mass.)

◀ Solve differential equation ▶

The constant acceleration formulas from Subsection 1.2 could be used since $a = g$ is constant, but we choose a more general method that applies even when a is not constant.

There are several approaches to obtaining the equations needed to solve the problem. One approach requires that we first replace a by dv/dt to obtain

$$\frac{dv}{dt} = g.$$

This is integrated to give

$$v = gt + A,$$

where A is a constant.

The initial condition that the marble is initially at rest (i.e. $v = 0$ when $t = 0$) gives $A = 0$, so

$$v = gt. \tag{12}$$

Now replacing v by dx/dt gives

$$\frac{dx}{dt} = gt.$$

This is integrated to obtain

$$x = \tfrac{1}{2}gt^2 + B,$$

where B is a constant.

The origin was chosen so that $x = 0$ when $t = 0$, so $B = 0$, which gives

$$x = \tfrac{1}{2}gt^2. \tag{13}$$

◀ Interpret solution ▶

When the marble hits the water, $x = 77$ and equation (13) gives $77 = \frac{1}{2}gt^2$, which on putting $g = 9.81\,\mathrm{m\,s^{-2}}$ gives

$$t = \sqrt{\frac{2 \times 77}{9.81}} \simeq 3.962.$$

So the model predicts that the marble hits the water approximately 3.96 seconds after being released.

(b) Putting $t = 3.962$ into equation (12) gives

$$v = 9.81 \times 3.962 \simeq 38.87.$$

So the model predicts that the marble has a speed of approximately 38.9 metres per second just before it hits the water.

Try the following exercise, following the same steps as given in the margin in the above example.

Exercise 14

A stone, dropped from rest, takes 3 seconds to reach the bottom of a well. Assume that the only force acting on the stone is gravity.

(a) Estimate the depth of the well.

(b) Estimate the speed of the stone when it reaches the bottom.

Use the more general approach of Example 4 rather than the constant acceleration formulas of Subsection 1.2.

The steps highlighted in Example 4 and in the solution to Exercise 14 are re-stated in the following procedure for solving mechanics problems involving one-dimensional motion using Newton's second law. Remember that it is intended to be a guide rather than a rigid set of rules.

Procedure 1 Applying Newton's second law in one dimension

Given a mechanics problem involving one-dimensional motion in which a question regarding the motion is to be answered, proceed as follows.

1. Draw a sketch of the physical situation, and annotate it with any relevant information. ◀ Draw picture ▶
2. Choose the x-axis to lie along the direction of motion, and select an origin. Mark the x-axis, its direction and the origin on your sketch. ◀ Choose axes ▶
3. State any assumptions that you make about the object and the forces acting on it. ◀ State assumptions ▶
4. Draw a force diagram. ◀ Draw force diagram ▶
5. Apply Newton's second law to obtain a vector equation. Resolve each force along the chosen axis in order to resolve the vector equation into a scalar equation. ◀ Apply Newton's 2nd law ▶
6. Substitute $v\,dv/dx$, dv/dt or d^2x/dt^2 for the acceleration a in the equation of motion, as appropriate, and solve the resulting differential equation(s) to obtain velocity v in terms of position x or time t, or position x in terms of time t, as required. ◀ Solve differential equation ▶
7. Interpret the solution in terms of the original problem. ◀ Interpret solution ▶

We make the following notes about this procedure.

- The procedure assumes that the question to be answered is given. When modelling real-world situations, deriving a suitable question is half the work. This part of the modelling process is looked at in detail in Unit 8.
- State any assumptions that you make (e.g. that the object is modelled as a particle and the only force acting on it is its weight).
- The importance of drawing a picture cannot be stressed too strongly. Include in the picture all relevant information from the problem, e.g. distances and masses.
- The choices of the origin and the direction of the x-axis are arbitrary, and will have no effect on the final outcome of your calculations. However, try to make these choices so that the position x and/or the velocity v are positive for the particle's motion, as this will simplify the algebra. If there is a clear starting point for the motion, then this is often a suitable choice for the origin, provided that it is a fixed point.

See Subsection 4.1.

- For some problems it may be necessary to choose a y-axis and even a z-axis as well as an x-axis.

You experienced the interchangeability of these steps in Unit 2.

- Steps 2 and 4 are interchangeable; the force diagram does not change with a different choice of axes. If you have difficulty choosing an x-axis, then draw the force diagram first.
- Your choice of substitution for a will depend on what question you want to answer (e.g. you may want an equation linking velocity to time, or velocity to position). (See Exercise 7 for examples of this.)
- When the acceleration is constant, you can use the general constant acceleration formulas from Subsection 1.2 (with a reference to any formula that you use) to solve the differential equation.
- Perform any readily available checks on your working, and consider whether your answers are physically reasonable. For example, you could check that the units of your answer are correct, or use common sense to tell you whether your answer is in the correct range.
- When you have finished interpreting the solution, write out your conclusion in words, and remember to include the physical units for any quantities given (as in Example 4). Also look back at the problem and check that you have fully answered the question asked.

Try using Procedure 1 in the following exercise.

Exercise 15

A ball is thrown vertically upwards from ground level with an initial speed of 10 metres per second. Assume that gravity is the only force acting on the ball.

(a) Find the time taken for the ball to reach its maximum height.

(b) Find the maximum height attained.

(c) Find the time taken for the ball to return to the ground.

(d) Find the speed of the ball as it reaches the ground on its return.

Exercise 16

A man leaning out from a window throws a ball vertically upwards from a point 4.4 metres above the ground. The initial speed of the ball is 7.6 metres per second. It travels up and then down in a straight vertical line, and eventually reaches the ground. Assume that its weight is the only force acting on the ball.

(a) Estimate the time that elapses before the ball reaches the ground.

(b) Estimate the speed of the ball when it strikes the ground.

In Example 4, which considered the motion of a marble falling from the Clifton Suspension Bridge, first the speed and then the position were found as functions of time (by using the substitutions dv/dt for a and dx/dt for v). However, as was indicated in the solution to that example, there are other approaches to solving the problem. You are asked to adopt one of these other approaches in the following exercise.

Exercise 17

A marble, initially at rest, is dropped from the Clifton Suspension Bridge and falls into the River Avon, 77 metres below. Assume that the only force acting on the marble is its weight due to gravity.

(a) By putting $a = v\,dv/dx$ in Newton's second law and solving the resulting differential equation, find the marble's speed v as a function of the distance x through which it has fallen.

(b) By putting $v = dx/dt$ in your answer to part (a), find the time t that the object takes to fall a distance x.

(c) Hence find the time taken before the marble hits the water, and the speed of the marble just before it hits the water.

In the previous examples and exercises, all of the forces acting on an object were in the same direction; but this does not have to be so for the motion to be in one dimension. This is illustrated in the following example.

Example 5

A crate of empty bottles of total mass 30 kilograms is being hauled by rope up a smooth ramp from the cellar of a pub. The ramp makes an angle of $\pi/6$ radians with the horizontal. When the crate has been hauled 2 metres up the ramp (i.e. 2 metres along the slope of the ramp), the rope suddenly breaks. It is estimated that if the crate hits the bottom of the ramp at a speed of 5 metres per second or greater, then the bottles in the crate will break.

A similar problem was considered in Unit 2. The problem is reconsidered in the presence of friction between the crate and the ramp in Section 4.

Assuming that the crate can be modelled as a particle, that there is no friction between the crate and the ramp, and that air resistance can be neglected, will the bottles break?

Solution

◀ Draw picture ▶

The situation is sketched in Figure 13.

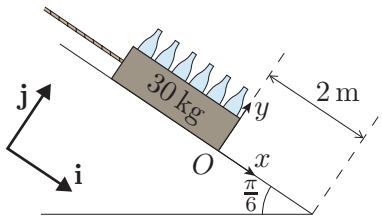


Figure 13 The crate of bottles on the ramp

◀ Choose axes ▶

The crate moves down the slope, so we choose the x-axis to point down the slope, with the origin O at the crate's position when the rope breaks. (We model the crate as a particle, so its position is at a point.) Choose a y-axis perpendicular to this, as shown in Figure 13. Also shown in the picture are the unit vectors $\mathbf{i}$ and $\mathbf{j}$ implied by this choice of axes.

◀ State assumptions ▶

The only forces on the crate are the weight $\mathbf{W}$ and the normal reaction $\mathbf{N}$ of the ramp (since we are neglecting friction and air resistance in this example). So modelling the crate as a particle of mass m, the force diagram (showing also how the relevant angles are calculated) is as shown in Figure 14.

◀ Draw force diagram ▶

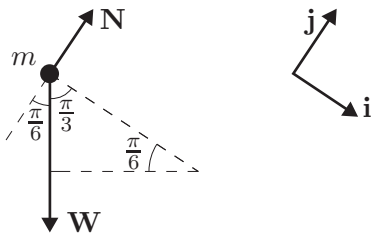


Figure 14 The force diagram

◀ Apply Newton's 2nd law ▶

Now, Newton's second law for this system gives

$$m\mathbf{a} = \mathbf{W} + \mathbf{N}. \tag{14}$$

As we have done elsewhere, we use the letter m rather than the value 30 kg for the mass since, as will be discussed in Unit 8, inserting actual data values too early in the modelling process can obscure features of the resulting model.

From the force diagram, $\mathbf{N} = |\mathbf{N}|\,\mathbf{j}$. The weight can be resolved into components using the formula from Unit 2, Example 9:

$$\mathbf{W} = |\mathbf{W}|\cos\tfrac{\pi}{3}\,\mathbf{i} - |\mathbf{W}|\sin\tfrac{\pi}{3}\,\mathbf{j} = \tfrac{1}{2}|\mathbf{W}|\,\mathbf{i} - \tfrac{\sqrt{3}}{2}|\mathbf{W}|\,\mathbf{j}.$$

The motion is along the slope, so $\mathbf{a} = a\mathbf{i}$. (All we are saying here is that there is no resultant force in the $\mathbf{j}$-direction.) Now that we have all the vectors in component form, we can immediately resolve equation (14) in the $\mathbf{i}$-direction:

$$ma = \tfrac{1}{2}|\mathbf{W}| + 0.$$

Since there is no motion in the $\mathbf{j}$-direction, we do not need to resolve in this direction. If we did, we would have

$$|\mathbf{N}| - mg\tfrac{\sqrt{3}}{2} = 0.$$

Substituting $|\mathbf{W}| = mg$ gives $ma = \frac{1}{2}mg$, so

$a = \frac{1}{2}g$.

Again m cancels out, so the results apply to a crate of any mass.

◀ Solve differential equation ▶

We want an equation for the velocity v in terms of the distance travelled x. Since the acceleration is constant, we can use equation (8), that is,

$$v^2 = v_0^2 + 2ax.$$

This gives the final velocity in terms of known quantities: the initial velocity v_0 is $0\,\mathrm{m\,s^{-1}}$, the distance travelled x is 2 m, and the constant acceleration is $a = \frac{1}{2}g$.

◀ Interpret solution ▶

Putting the data into the equation gives

$$v^2 = 0^2 + 2 \times \tfrac{1}{2}g \times 2 \simeq 19.62 \quad \text{to 2 d.p.}$$

So $v \simeq 4.4$ to one decimal place, which means that the crate is travelling at about 4.4 metres per second when it hits the bottom of the ramp. This speed is just below the estimated speed at which the bottles will break, so – provided that the estimate of the speed at which the bottles will break is a good one – this model predicts that the bottles will not break.

Exercise 18

A skier of mass 65 kilograms starts from rest at the top of a 120-metre ski slope (i.e. 120 metres is the vertical distance from top to bottom), as shown in Figure 15.

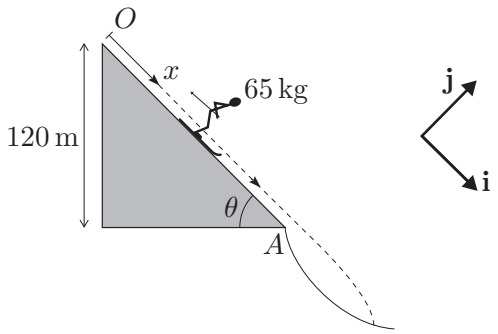


Figure 15 The skier on the ski slope

The skier travels down the slope before leaving the ground and then landing some distance further down the hill. Use the x-axis shown, which points down the slope with the origin O at the top of the slope and point A at the bottom of the slope. Let θ be the angle that the slope makes with the horizontal. Assume that friction and air resistance can be neglected.

This problem is reconsidered in the presence of friction in Section 4.

(a) If $\theta = \frac{\pi}{4}$, what is the speed of the skier at point A at the bottom of the slope?

(b) If $\theta = \frac{\pi}{3}$, what is the speed of the skier at point A at the bottom of the slope?

4 Some more force models

In Section 3 we described a procedure for predicting the motion of a particle using Newton's second law. To use the procedure, we must first model every force acting on the particle. In this section, models of forces due to friction and air resistance are described. This greatly extends the range of problems that can be solved using the procedure.

4.1 Friction

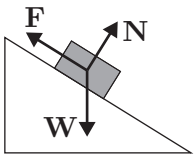


Figure 16 The forces acting on a block at rest on a slope

In Unit 2 you encountered a model for the friction force that acts on an object at rest in contact with a surface. This force $\mathbf{F}$ has magnitude less than or equal to $\mu|\mathbf{N}|$, that is, $|\mathbf{F}| \leq \mu|\mathbf{N}|$, where $\mathbf{N}$ is the normal reaction force and μ is the coefficient of static friction between the object and the surface. The direction of the friction force opposes any possible motion (e.g. is up the slope in Figure 16).

The coefficient μ is called the coefficient of *static* friction to distinguish it from the **coefficient of sliding friction**, denoted by μ', that is used when an object is moving along a surface. In this situation, experiments show that the magnitude of the friction force $\mathbf{F}$ is equal to $\mu'|\mathbf{N}|$, that is, $|\mathbf{F}| = \mu'|\mathbf{N}|$, where $\mathbf{N}$ is the normal reaction force and μ' is the coefficient of sliding friction; the direction of the friction force is opposite to the direction of motion.

The names *coefficient of dynamical friction* or *coefficient of kinetic friction* are sometimes used instead of the name *coefficient of sliding friction*.

Modelling sliding friction

- The friction force $\mathbf{F}$ acts in a direction perpendicular to the normal reaction $\mathbf{N}$ and opposite to the motion.
- $|\mathbf{F}| = \mu'|\mathbf{N}|$, where μ' is the coefficient of sliding friction for the two surfaces involved.

The numerical value of the coefficient of sliding friction μ' is always smaller than the numerical value of the coefficient of static friction μ. (It is harder to get objects moving than to keep them moving.) A generalisation that is often useful is

$$\mu' \simeq \tfrac{3}{4}\mu.$$

The use of the sliding friction model is best explained by an example.

Example 6

Consider again the problem of Example 5 concerning the crate of empty bottles sliding down a cellar ramp, but this time assume that there is friction between the crate and the ramp, with coefficient of sliding friction $\mu' = 0.15$.

(a) Estimate the speed of the crate when it reaches the bottom of the ramp. Will the bottles break?

(b) Compare the answer to part (a) with the answer to Example 5, and comment.

In Example 5 it was assumed the bottles would break if the speed at the bottom of the ramp exceeded $5\,\mathrm{m\,s^{-1}}$.

Solution

(a) The situation is sketched in Figure 17.

◀ Draw picture ▶

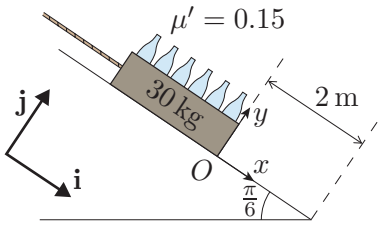


Figure 17 The crate of bottles on the ramp

Choose the x-axis parallel to the ramp and the y-axis perpendicular to it, as shown in Figure 17. Choose the origin O to be the crate's position when the rope breaks.

◀ Choose axes ▶

The crate is modelled as a particle. The forces acting on the crate are its weight $\mathbf{W}$, the normal reaction of the ramp on the crate $\mathbf{N}$, and the friction force between the ramp and the crate $\mathbf{F}$, giving the force diagram in Figure 18.

◀ State assumptions ▶

◀ Draw force diagram ▶

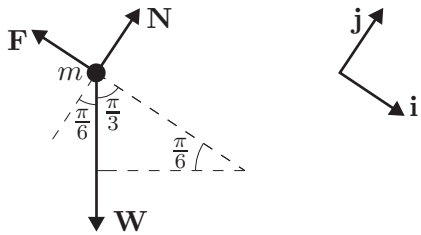


Figure 18 The force diagram

Using m for the mass of the crate (as in Example 5), applying Newton's second law to this system gives

◀ Apply Newton's 2nd law ▶

$$m\mathbf{a} = \mathbf{W} + \mathbf{N} + \mathbf{F}. \tag{15}$$

From the force diagram, $\mathbf{N} = |\mathbf{N}|\,\mathbf{j}$ and $\mathbf{F} = -|\mathbf{F}|\,\mathbf{i}$. It is also apparent that the motion is one-dimensional parallel to the x-axis, that is, $\mathbf{a} = a\mathbf{i}$. The weight is resolved into components as

$$\begin{aligned}\mathbf{W} &= |\mathbf{W}|\cos\tfrac{\pi}{3}\,\mathbf{i} - |\mathbf{W}|\sin\tfrac{\pi}{3}\,\mathbf{j}\\ &= \tfrac{1}{2}mg\,\mathbf{i} - \tfrac{\sqrt{3}}{2}mg\,\mathbf{j}.\end{aligned}$$

Now we can immediately resolve equation (15) in the $\mathbf{i}$-direction to obtain

$$ma = \tfrac{1}{2}mg + 0 - |\mathbf{F}|.$$

Using the friction model $|\mathbf{F}| = \mu'|\mathbf{N}|$ gives

$$ma = \tfrac{1}{2}mg - \mu'|\mathbf{N}|. \tag{16}$$

So to find a, we need to find $|\mathbf{N}|$. To do this, we resolve equation (15) in the $\mathbf{j}$-direction to obtain

$$0 = -\tfrac{\sqrt{3}}{2}mg + |\mathbf{N}| + 0,$$

which gives $|\mathbf{N}| = \frac{\sqrt{3}}{2}mg$. Substituting this into equation (16) gives

$$ma = \tfrac{1}{2}mg - \tfrac{\sqrt{3}}{2}\mu' mg,$$

so

$$a = \tfrac{1}{2}g - \tfrac{\sqrt{3}}{2}\mu' g.$$

◀ Solve differential equation ▶

As in Example 5, the acceleration is constant, so we can use

$$v^2 = v_0^2 + 2ax.$$

This gives the final speed in terms of known quantities: the initial velocity $v_0 = 0$, the distance travelled $x = 2$, and the constant acceleration $a = \frac{1}{2}g - \frac{\sqrt{3}}{2}\mu' g$, where $\mu' = 0.15$.

◀ Interpret solution ▶

Putting the data into the equation gives

$$v^2 = 0^2 + 2 \times (\tfrac{1}{2}g - \tfrac{\sqrt{3}}{2} \times 0.15 \times g) \times 2 \simeq 14.52 \quad \text{to 2 d.p.}$$

So $v \simeq 3.8$ to one decimal place, which means that the crate is travelling at about 3.8 metres per second when it hits the bottom of the ramp. This speed is well below the estimated speed at which the bottles will break, so this model predicts that the bottles will not break.

(b) The model that neglected friction (Example 5) predicted a speed of $4.4\,\mathrm{m\,s^{-1}}$. The new prediction of $3.8\,\mathrm{m\,s^{-1}}$ is slower, as should be expected after the incorporation of friction into the model (this is a good quick check of the solution), and is significantly slower than the breaking threshold of $5\,\mathrm{m\,s^{-1}}$. So in this case, even if the estimated breaking threshold is not very accurate, we can be reasonably confident that our prediction that the bottles will not break is correct.

Table 3 shows some values of the coefficient of sliding friction that may be useful in problems involving sliding objects.

Table 3 Coefficients of sliding friction

Object	Surface	μ'
Waxed ski	Dry snow	0.03
Brass	Ice	0.02
Vulcanised rubber	Dry tarmac	1.07
Vulcanised rubber	Wet tarmac	0.95

Exercise 19

A tip-up truck is delivering a concrete block to a building site. The driver increases the angle of tip of the carrier until the concrete block begins to slide, then keeps the carrier at this constant angle. The coefficient of static friction between a concrete block and metal is approximately $\mu = 0.4$, and the coefficient of sliding friction is approximately $\mu' = 0.3$.

(a) Calculate the angle at which the concrete block begins to slide.

(b) If the concrete block has 3 metres to travel before leaving the carrier, how long will it take to unload it?

Exercise 20

Repeat Exercise 18 under the new modelling assumption that friction cannot be neglected; the coefficient of sliding friction between waxed skis and dry snow is 0.03. Compare your answers here with those for Exercise 18, and comment.

4.2 Air resistance

In many situations it is adequate to treat a falling object taking into account only the force of gravity. For example, we predicted a time of fall of a marble from the Clifton Suspension Bridge of 3.96 seconds, which is close to an experimental value of 4.1 seconds. Most of the time difference can be attributed to air resistance, which is modelled in this subsection.

The idea of air resistance is quite familiar in everyday life. For example, any experienced cyclist knows the following.

- Air resistance tends to slow one down and resists one's attempts to increase speed.
- At low speeds air resistance has little effect, but at higher speeds it becomes more noticeable, making it difficult to cycle faster than about 40 kilometres per hour.
- Air resistance can be reduced by crouching over the handlebars to present a smaller profile to the wind.

From these observations we conclude that air resistance is a force $\mathbf{R}$ whose direction is opposite to that of the motion of an object and whose magnitude $|\mathbf{R}|$ depends on the object's speed, shape and size.

To simplify the discussion, we will restrict our attention to smooth spherical objects. For such an object, we would expect $|\mathbf{R}|$ to increase as the object's speed $|\mathbf{v}|$ and diameter D increase. In fact, experiments show that the force on a sphere depends only on the product $D|\mathbf{v}|$. The results of such experiments are shown as the solid black line in Figure 19 (where, because of the wide range of values, a log–log graph has been used).

Notice the 'kink' in Figure 19 when $\log_{10}(D|\mathbf{v}|)$ is between 0 and 1: in this region there is less air resistance as the sphere travels faster! The physical reason for this is that a highly chaotic layer of air develops around the sphere.

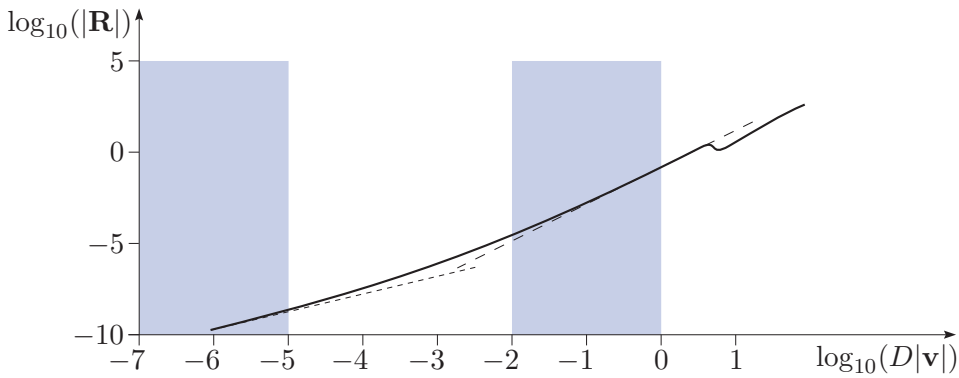


Figure 19 The experimental data for air resistance together with broken lines showing the two simplest models and their ranges of validity (shaded)

The graph has a complicated shape, so the air resistance force is modelled by a complicated function of $D|\mathbf{v}|$. However, there are two simple models that fit the experimental data for wide ranges of situations, namely the linear and quadratic models

$$|\mathbf{R}| = c_1 D|\mathbf{v}| \quad \text{and} \quad |\mathbf{R}| = c_2 D^2|\mathbf{v}|^2,$$

where c_1 and c_2 are positive constants. The best fit of these models to the experimental data (when using SI units) is given by $c_1 = 1.7 \times 10^{-4}$ and $c_2 = 0.20$ over certain ranges:

$$|\mathbf{R}| \simeq 1.7 \times 10^{-4} D|\mathbf{v}| \quad \text{for } D|\mathbf{v}| \lesssim 10^{-5};$$
$$|\mathbf{R}| \simeq 0.2 D^2|\mathbf{v}|^2 \quad \text{for } 10^{-2} \lesssim D|\mathbf{v}| \lesssim 1.$$

The symbol $\lesssim$ means 'less than about'.

These approximations for air resistance are shown as broken lines in Figure 19, together with their ranges of validity (shaded).

Consider a sphere of given diameter D. From Figure 19 it can be seen that the linear model applies for low velocities, and the quadratic model applies for an intermediate range of velocities. Notice also that there are large ranges of velocities where neither of these simple models applies.

The air resistance force $\mathbf{R}$ always opposes the motion, that is, it is in the direction of $-\mathbf{v}$. So the vector equations for the air resistance models given above are

$$\mathbf{R} = -c_1 D\mathbf{v} \quad \text{and} \quad \mathbf{R} = -c_2 D^2|\mathbf{v}|\mathbf{v}.$$

Note that the vector and scalar statements of the quadratic air resistance model agree since

$$\begin{aligned} |-c_2 D^2|\mathbf{v}|\mathbf{v}| &= c_2 D^2|\mathbf{v}|\,|\mathbf{v}| \\ &= c_2 D^2|\mathbf{v}|^2. \end{aligned}$$

Air resistance

The air resistance force $\mathbf{R}$ on a smooth spherical object of diameter D travelling with velocity $\mathbf{v}$ can be modelled as follows:

$$\text{linear model} \quad \mathbf{R} = -c_1 D\mathbf{v} \quad \text{for } D|\mathbf{v}| \lesssim 10^{-5}, \tag{17}$$
$$\text{quadratic model} \quad \mathbf{R} = -c_2 D^2|\mathbf{v}|\mathbf{v} \quad \text{for } 10^{-2} \lesssim D|\mathbf{v}| \lesssim 1, \tag{18}$$

where $c_1 \simeq 1.7 \times 10^{-4}$ and $c_2 \simeq 0.20$.

These models can be also used for other fluids, with different ranges of applicability. For example, in water we have $c_1 = 9.4 \times 10^{-3}$ for $D|\mathbf{v}| \lesssim 10^{-6}$ and $c_2 = 156$ for $10^{-3} \lesssim D|\mathbf{v}| \lesssim 10^{-1}$.

The use of these models is illustrated by examples, the first of which adds air resistance to Example 4.

Example 7

In Example 4, the falling time for a marble dropped from the Clifton Suspension Bridge into the River Avon was calculated to be 3.96 seconds. The experimental value is 4.1 seconds. Can the discrepancy be accounted for by a linear air resistance model?

The investigation is subdivided into the following steps.

(a) Find how, under a linear air resistance model, the distance from the point of release varies with time for an arbitrary spherical object of mass m and diameter D.

(b) Calculate the time, under a linear air resistance model, that a marble of diameter 2 cm and mass 13 g takes to fall the 77 m from the Clifton Suspension Bridge into the River Avon below.

(c) Comment on the validity of the linear air resistance model for this problem.

Solution

(a) The picture is the same as in Example 4 – see Figure 11. ◀ Draw picture ▶

The x-axis points vertically downwards, with its origin O at the point where the marble is released. The unit vector $\mathbf{i}$ acts downwards. ◀ Choose axes ▶

There are two forces acting on the marble, namely its weight due to gravity downwards and air resistance upwards, as shown in the force diagram in Figure 20. ◀ State assumptions ▶ ◀ Draw force diagram ▶

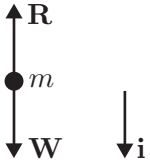


Figure 20 The force diagram

Applying Newton's second law to the marble gives ◀ Apply Newton's 2nd law ▶

$$m\mathbf{a} = \mathbf{W} + \mathbf{R}. \tag{19}$$

From the force diagram, $\mathbf{W} = |\mathbf{W}|\,\mathbf{i} = mg\mathbf{i}$. Using the linear air resistance model, we have $\mathbf{R} = -c_1 D\mathbf{v} = -c_1 Dv\mathbf{i}$, since $\mathbf{v} = v\mathbf{i}$ with $v \geq 0$. The acceleration is also downwards, so $\mathbf{a} = a\mathbf{i}$. Now we can resolve equation (19) in the $\mathbf{i}$-direction to obtain

$$ma = mg - c_1 Dv.$$

◀ Solve differential equation ▶

We want the distance x in terms of the time t. One way of obtaining this is to first use the substitution $a = dv/dt$, and then substitute dx/dt for v later. So we have

$$m\frac{dv}{dt} = mg - c_1 Dv.$$

Notice that the acceleration is not constant here, so we cannot use the constant acceleration formulas.

This assumption is checked later. It is certainly true initially as the marble is dropped from rest.

We now make the assumption that $mg - c_1 Dv > 0$ (i.e. the air resistance never becomes so strong as to overcome the marble's weight). Under this assumption, the differential equation can be solved by the separation of variables method, to obtain

$$\int \frac{m}{mg - c_1 Dv}\, dv = \int 1\, dt,$$

The integrating factor method could also be used.

which, on dividing the numerator and denominator by m, and for convenience writing $k = C_1 D/m$, gives

$$\int \frac{1}{g - kv}\, dv = \int 1\, dt.$$

Note that $g - kv > 0$ if $mg - c_1 Dv > 0$.

Integrating, using the assumption that $g - kv > 0$, leads to

$$-\frac{1}{k}\ln(g - kv) = t + A, \tag{20}$$

where A is an arbitrary constant. Rearranging gives

$$g - kv = e^{(-kt - kA)} = Be^{-kt},$$

where $B = e^{-kA}$ is another constant. The initial condition that the marble is initially at rest (i.e. $v = 0$ when $t = 0$) gives $B = g$, so

$$v = \frac{g}{k}(1 - e^{-kt}). \tag{21}$$

Since c_1, D, t and m are all greater than or equal to zero, so that $kt > 0$, this equation predicts that $v < g/k = mg/c_1 D$ for all t, which is consistent with the assumption made in obtaining equation (20).

Writing dx/dt for v gives

$$\frac{dx}{dt} = \frac{g}{k}(1 - e^{-kt}).$$

Integrating directly gives

$$x = \frac{g}{k}t + \frac{g}{k^2}e^{-kt} + C,$$

where C is a constant.

The initial condition $x = 0$ when $t = 0$ gives $C = -g/k^2$, so

$$x = \frac{g}{k}t - \frac{g}{k^2}(1 - e^{-kt}), \tag{22}$$

where $k = c_1 D/m$.

Since $e^{-kt} \to 0$ as $t \to \infty$, the speed v tends to g/k, and the position x tends to $gt/k - g/k^2 \to gt/k$.

◀ Interpret solution ▶

(b) We want to find t given $x = 77$. To do this, we could try to rearrange equation (22) to give t in terms of x; however, no such rearrangement is possible. But we can use a numerical method. First we use the given data to interpret equation (22) in the context of the current problem.

We have

$$k = \frac{c_1 D}{m} = \frac{1.7 \times 10^{-4} \times 0.02}{0.013} \simeq 2.6 \times 10^{-4},$$

so

$$x \simeq \frac{9.81}{2.6 \times 10^{-4}}\, t - \frac{9.81}{(2.6 \times 10^{-4})^2}\left(1 - e^{-2.6 \times 10^{-4} \times t}\right). \qquad (23)$$

Now we already have two values for the time taken for the marble to fall 77 metres: 3.96 seconds estimated in Example 4, and 4.1 seconds given by experiment. Substituting these values for t into equation (23), we obtain

$$x(3.96) \simeq 76.89 \quad \text{and} \quad x(4.1) \simeq 82.42.$$

So it looks like $t = 3.96$ is close to the solution of equation (23) for $x = 77$.

So $t = 3.96$ (to two decimal places) when $x = 77$, that is, the marble takes about 3.96 seconds to reach the River Avon under the linear air resistance model. This is exactly the same value, to two decimal places, as for the model without air resistance (Example 4).

More detailed calculation shows that the linear air resistance model predicts that the marble will hit the water approximately one thousandth of a second later than it will in the case of the model without air resistance.

(c) A condition for the linear air resistance model to be valid is that the product of the diameter of the marble and its speed is less than about 10^{-5}. Using equation (21) gives $|\mathbf{v}(3.96)| \simeq 38.8$, so $D|\mathbf{v}| \simeq 0.78$, which is much greater than $10^{-5}\,\mathrm{m\,s^{-1}}$. So the linear air resistance model is not appropriate.

In fact, $D|\mathbf{v}(t)| \lesssim 10^{-5}$ for $t \lesssim 5.1 \times 10^{-5}$, so the linear model applies for only about the first 51 microseconds of the motion of the marble.

From Example 7 it seems that the linear air resistance model may apply only to objects moving very slowly – not to the speeds experienced in everyday life. So our attention turns to the quadratic air resistance model. In general, the differential equations that arise from this model are harder to solve than the ones that arise from the linear model. However, some of the differential equations are easily soluble, as the following example shows.

Example 8

Revisit the Clifton Suspension Bridge problem in Example 7 using the quadratic air resistance model.

(a) Derive an expression for the marble's velocity in terms of its position.

(b) Use the expression derived in part (a) to estimate the speed of the marble just before it hits the water.

(c) Is the quadratic air resistance model valid for this problem?

Solution

(a) Everything is the same as in Example 7 up to the point where we apply Newton's second law to the marble, to obtain

◀ Apply Newton's 2nd law ▶

$$m\mathbf{a} = \mathbf{W} + \mathbf{R}. \qquad (24)$$

As before, $\mathbf{W} = mg\mathbf{i}$, but now we use the quadratic air resistance model $\mathbf{R} = -c_2D^2|\mathbf{v}|\mathbf{v}$. Since the motion is downwards, $\mathbf{v} = v\mathbf{i}$ with $v > 0$ and hence $|\mathbf{v}| = v$, so $\mathbf{R} = -c_2D^2v^2\mathbf{i}$. Resolving equation (24) in the $\mathbf{i}$-direction gives

$$ma = mg - c_2D^2v^2.$$

◀ Solve differential equation ▶

The question asks for v in terms of x, so we use the substitution $a = v\,dv/dx$ to obtain

$$mv\frac{dv}{dx} = mg - c_2D^2v^2.$$

This equation is non-linear in v thus cannot be solved by the integrating factor method.

We now make the assumption that $mg - c_2D^2v^2 > 0$ (i.e. the air resistance never becomes so strong as to overcome the marble's weight). Under this assumption, the differential equation can be solved by separation of variables,

This assumption is checked later.

$$\int \frac{mv}{mg - c_2D^2v^2}\,dv = \int 1\,dx,$$

which is equivalent to

$$\int \frac{v}{g - kv^2}\,dv = \int 1\,dx,$$

where $k = c_2D^2/m$. Integrating, using the assumption that $g - kv^2 > 0$, and noting that

Note that $g - kv^2 > 0$ if $mg - c_2D^2v^2 > 0$, as assumed. The case where $g - kv^2 < 0$ is considered later.

$$\frac{d}{dv}(g - kv^2) = -2kv,$$

so the integral is of the form

$$-\frac{1}{2k}\int \frac{f'(v)}{f(v)}\,dv,$$

we have

$$-\frac{1}{2k}\ln(g - kv^2) = x + A,$$

where A is an arbitrary constant. So

$$g - kv^2 = e^{(-2kx-2kA)} = Be^{-2kx},$$

where $B = e^{-2kA}$ is another constant. The initial condition that the marble is initially at rest at the origin (i.e. $v = 0$ when $x = 0$) gives $B = g$, so

$$g - kv^2 = ge^{-2kx}, \quad \text{or equivalently,} \quad v^2 = \frac{g}{k}(1 - e^{-2kx}).$$

Therefore

$$v = \sqrt{\frac{g}{k}}\sqrt{1 - e^{-2kx}}, \tag{25}$$

where $k = c_2D^2/m$ and we take the positive square root since we must have $v \geq 0$ from the description of the problem.

Note that c_2, D, x and m are all greater than or equal to zero, so $0 < \exp(-2kx) \leq 1$, and equation (25) predicts that $v < \sqrt{g/k} = \sqrt{mg/c_2D^2}$ for all t, which is consistent with the assumption made earlier.

Since $e^{-2kx} \to 0$ as $x \to \infty$, the speed v tends to $\sqrt{g/k}$.

(b) Substituting the values for the marble (see Example 7) gives

◀ Interpret solution ▶

$$k = \frac{c_2 D^2}{m} = \frac{0.2 \times (0.02)^2}{0.013} \simeq 6.2 \times 10^{-3},$$

so from equation (25),

$$v = \sqrt{\frac{9.81}{6.2 \times 10^{-3}}}\sqrt{1 - e^{-2 \times 6.2 \times 10^{-3} \times 77}} \simeq 31.2.$$

So the model predicts that the speed of the marble is about 31 metres per second just before it hits the water.

(c) To test whether the quadratic air resistance model applies to the motion, we must calculate $D|\mathbf{v}|$ and check that it is in the range 0.01 to 1. At the end of the motion, $D|\mathbf{v}(77)| \simeq 0.62$, which is within the limits of applicability of the model. The motion starts from rest, and $D|\mathbf{v}(0)| = 0$ is obviously outside the limits of the model. However, $D|\mathbf{v}(0.01)| \simeq 0.01$, so after the marble has dropped approximately one centimetre, the quadratic model applies. So the quadratic model applies for almost all of the motion.

Admittedly, this calculation has been done using the quadratic model, but it can't be far out!

The given air resistance models apply only to smooth spherical objects. To apply these models to other objects, we have to model the objects as smooth spheres. The diameter of the sphere used to model an object is referred to as the **effective diameter** of the object. The following exercise makes use of this idea.

The determination of an effective diameter for an object is beyond the scope of this module.

Exercise 21

In the discussion leading up to the statement of Newton's second law in Section 2, the motion of an empty toboggan sliding on ice was considered. In this exercise, the effect of air resistance on the motion is examined.

Assume that the quadratic air resistance model applies, and that the effective diameter of the toboggan is 5 centimetres. If the toboggan has mass 2 kilograms and an initial speed of 2 metres per second, by what percentage has its speed diminished after it has travelled 100 metres?

Terminal speed

The general equation for the velocity $\mathbf{v} = v\mathbf{i}$ of an object falling from rest under gravity, ignoring the effects of air resistance, was derived in Example 4 as

$$v = gt.$$

This predicts that v increases indefinitely as t increases.

The general equation for the velocity $\mathbf{v} = v\mathbf{i}$ of an object falling from rest under gravity with linear air resistance was derived in Example 7 to be

$$v = \frac{g}{k}(1 - e^{-kt}),$$

where $k = c_1 D/m$. The exponential term decreases to zero as time increases, so

$$v \to \frac{g}{k} = \frac{mg}{c_1 D} \quad \text{as } t \to \infty.$$

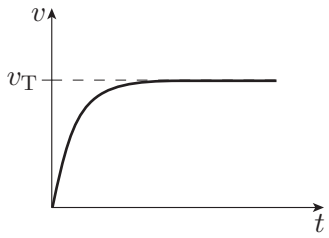


Figure 21 Velocity graph

This behaviour, which is quite different from the case of the model that neglects air resistance, is illustrated in Figure 21. The limiting value is called the **terminal speed** of the object and is denoted by v_T. So in the case of the linear air resistance model, the terminal speed (in SI units) is

$$v_T = \frac{g}{k} = \frac{mg}{c_1 D}.$$

The quadratic air resistance model predicts results that are qualitatively similar to the linear case. The general equation for the speed of an object falling from rest under gravity with quadratic air resistance was derived in Example 8 to be

$$v = \sqrt{\frac{g}{k}}\sqrt{1 - e^{-2kx}},$$

where $k = c_2 D^2/m$. The exponential term decreases to zero as x increases, so the quadratic model predicts a terminal speed (in SI units) given by

$$v_T = \sqrt{\frac{g}{k}} = \sqrt{\frac{mg}{c_2 D^2}}.$$

There is another way of looking at the concept of terminal speed. The equation of motion of an object falling from rest under gravity with linear air resistance is $ma = mg - c_1 Dv$; thus $v_T = mg/c_1 D$ is exactly that value of v for which $a = 0$. The equation of motion of an object falling from rest under gravity with quadratic air resistance is $ma = mg - c_2 D^2 v^2$; again $v_T = \sqrt{mg/c_2 D^2}$ is such that $a = 0$. Thus in each case the terminal speed is the speed at which the object can fall without accelerating, or in other words at which the air resistance just balances the object's weight. This way of looking at terminal speed enables us to derive equations for the terminal speed of any falling object, irrespective of whether or not the object is falling from rest, as the following example illustrates.

See Example 7.

See Example 8.

Example 9

A small spider is blown by the wind into the high atmosphere and, once the wind has died away, falls back to Earth. The spider may be modelled as a sphere of effective diameter 2 millimetres.

If the spider has mass 0.004 grams, calculate the speed at which it will land, assuming that the fall is long enough for the spider effectively to reach terminal speed.

Solution

The first step is to draw a picture and mark the x-axis on it, then draw the force diagram (see Figure 22).

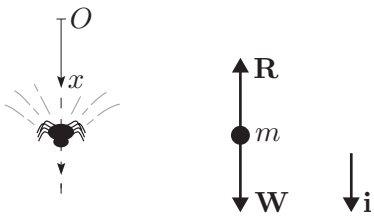


Figure 22 The spider descending to Earth, and the force diagram

When the spider is travelling at its terminal speed, it is not accelerating. Therefore by Newton's second law the resultant force on the spider must be zero, that is,

$$\mathbf{W} + \mathbf{R} = \mathbf{0}. \tag{26}$$

Suppose that the linear air resistance model is valid.

The force models are $\mathbf{W} = mg\mathbf{i}$ and $\mathbf{R} = -c_1 D\mathbf{v}_\mathrm{T}$. The motion is downwards, so $\mathbf{v}_\mathrm{T} = v_\mathrm{T}\mathbf{i}$. Resolving equation (26) in the $\mathbf{i}$-direction gives

$$mg - c_1 D v_\mathrm{T} = 0,$$

so

$$v_\mathrm{T} = \frac{mg}{c_1 D} = \frac{4 \times 10^{-6} \times 9.81}{1.7 \times 10^{-4} \times 0.002} \simeq 115.$$

So the terminal speed of the spider under the linear air resistance model is about 115 metres per second. This gives $D|\mathbf{v}_\mathrm{T}| \simeq 0.2$, which is much greater than 10^{-5}, so the linear air resistance model is not valid.

Now suppose that the quadratic air resistance model holds.

The weight is $\mathbf{W} = mg\mathbf{i}$, as before. The air resistance force is now given by the equation $\mathbf{R} = -c_2 D^2 |\mathbf{v}_\mathrm{T}| \mathbf{v}_\mathrm{T}$. As before, the motion is downwards, so $\mathbf{v}_\mathrm{T} = v_\mathrm{T}\mathbf{i}$ and $\mathbf{R} = -c_2 D^2 v_T^2 \mathbf{i}$. Resolving equation (26) in the $\mathbf{i}$-direction gives

$$mg - c_2 D^2 v_T^2 = 0,$$

which can be rearranged to obtain

$$v_\mathrm{T} = \sqrt{\frac{mg}{c_2 D^2}} = \sqrt{\frac{4 \times 10^{-6} \times 9.81}{0.2 \times (0.002)^2}} \simeq 7.$$

So the terminal speed of the spider under the quadratic air resistance model is about 7 metres per second. This gives $D|\mathbf{v}_T| \simeq 0.014$, which is in the range of validity for the quadratic model. So the quadratic air resistance model is valid and gives a terminal speed of about 7 metres per second.

The results about the terminal speed of objects falling under the influence of gravity and air resistance alone are summarised in the following box.

Terminal speed under air resistance

The terminal speed of an object falling under gravity and air resistance is the constant velocity that the object will acquire as time tends to infinity, and is the velocity at which air resistance just balances the object's weight. For an object of mass m and effective diameter D it is given (using SI units) as follows:

under the linear air resistance model $$v_{\mathrm{T}} = \frac{mg}{c_1 D}, \tag{27}$$

under the quadratic air resistance model $$v_{\mathrm{T}} = \sqrt{\frac{mg}{c_2 D^2}}, \tag{28}$$

where g is the magnitude of the acceleration due to gravity, $c_1 \simeq 1.7 \times 10^{-4}$ and $c_2 \simeq 0.20$.

The terminal speed of an object is an important concept that is often the only quantity of interest in a problem including air resistance. For example, it is often assumed that the landing speed of a parachute is essentially its terminal speed, so considering terminal speed is a crucial aspect of the design of parachutes.

Exercise 22

A parachutist of mass 65 kilograms has a parachute of effective diameter 10 metres when fully opened. Estimate the landing speed of the parachutist, assuming that the parachute jump is long enough for the terminal speed effectively to be reached.

Exercise 23

The maximum speed at which a parachutist can land safely is about 13 metres per second. Assuming that the parachute jump is long enough for the terminal speed effectively to be reached, calculate the effective diameter of a parachute that will enable a parachutist of mass 70 kilograms to land safely.

The parachutist will start in freefall before pulling the rip cord, thus will be approaching the terminal speed from above. It will be important that when the rip cord is pulled, there is sufficient time to slow down.

Exercise 24

Revisit the Clifton Suspension Bridge problem described in Example 8, with different initial conditions. Assume now that instead of it being dropped from rest, the marble is catapulted downwards with an initial velocity of $50\,\mathrm{m\,s^{-1}}$.

The initial stages of the solution will be exactly the same as in the example, since the initial conditions are not used until the step solving the differential equation. So the equation of motion of the marble is still

$$v\frac{dv}{dx} = g - kv^2,$$

where $k = c_2D^2/m$ is a positive constant.

(a) Derive an expression for the marble's velocity in terms of its position.

(b) Describe the motion of the marble.

---

5 Projectiles

In Section 1 we discussed motion in one dimension. In this section we discuss two-dimensional motion – more specifically, the motion of projectiles modelled as particles. We will use the following terminology in our discussion. During the period that the projectile is off the ground and subject only to the force of gravity (and possibly air resistance), it is said to be *in flight*. The start of the flight is the *launch*, and the initial velocity is the *launch velocity*. The flight ends with an *impact* (often hitting the ground again, but possibly hitting some other target). The *time of flight* is the time between the moment of launch and the moment of impact. The path of the projectile while in flight is its **trajectory**.

Whenever we mention a projectile in this section, it will be modelled as a particle.

In this section, we consider only models in which the effect of air resistance is assumed to be negligible.

In ignoring air resistance, we also ignore effects such as the swerve that may occur when a football is kicked with spin, or when a golf ball is sliced.

In Subsection 5.1, we model the forces acting on a projectile and solve its equation of motion. In Subsection 5.2, we look at the trajectory of a projectile and consider a variety of examples.

5.1 Motion of a projectile

We consider the motion of a particle, called a projectile, launched into the air. The only force acting on the projectile while it is in flight is that due to gravity – we are assuming that air resistance is negligible and can be ignored. So the subsequent motion is in two dimensions, that is, in the vertical plane determined by the direction of the launch velocity. We normally choose the origin to be at the point of launch, the x-axis to be in the direction of the launch velocity, and the y-axis to be vertically upwards. However, it is sometimes convenient to choose an origin that is not the point of projection. Figure 23 shows two examples of projectiles.

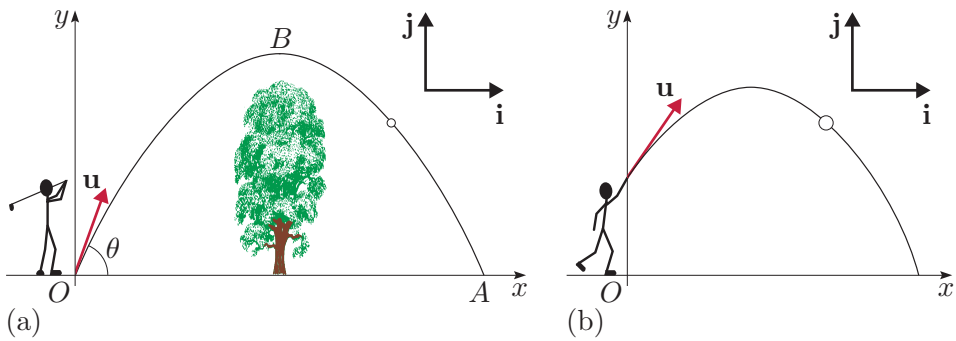


Figure 23 (a) Golf shot (b) Shot put

In Figure 23(a) the projectile is launched from the origin, while Figure 23(b) illustrates the case where the origin is not the launch position. Furthermore, it may be convenient to take $t = 0$ at some time other than the moment of projection. All these complications will change the initial conditions of the projectile motion. None changes the fundamental equation of motion, however. So long as we choose the y-axis (and the unit vector $\mathbf{j}$) to be vertically upwards, and ignore air resistance, the only force on a projectile of mass m is that due to gravity, $-mg\mathbf{j}$, and Newton's second law gives

$$m\,\ddot{\mathbf{r}}(t) = \mathbf{F} = -mg\mathbf{j},$$

that is,

$$\ddot{\mathbf{r}}(t) = -g\mathbf{j}. \tag{29}$$

We will integrate this equation twice, using the initial conditions, to obtain $\mathbf{r}(t)$.

In this subsection we look first at projectile motion of the type illustrated in Figure 23(a), where the launch point and the impact point are in the same horizontal plane. Assume that the projectile is launched with velocity $\mathbf{u}$, at an angle θ above the horizontal. We will derive expressions in terms of $\mathbf{u}$ and θ for the 'range' of the projectile's flight (OA in Figure 23(a)), and for the coordinates of the highest point of its trajectory (B in Figure 23(a)).

We refer to θ as the *launch angle*.

To solve the differential equation (29), we can integrate it twice. We will work in the vector form. Integrating once gives

$$\dot{\mathbf{r}}(t) = -gt\mathbf{j} + \mathbf{c}, \tag{30}$$

where $\mathbf{c}$ is a constant vector. Integrating equation (30) gives

$$\mathbf{r}(t) = -\tfrac{1}{2}gt^2\mathbf{j} + \mathbf{c}t + \mathbf{d}, \tag{31}$$

where $\mathbf{d}$ is a constant vector. We will take $t = 0$ to be the moment of launch and the origin to be the point of launch. Then we have the initial condition

$$\mathbf{r}(0) = \mathbf{0}.$$

Substituting this into equation (31) gives $\mathbf{d} = \mathbf{0}$, so we have

$$\mathbf{r}(t) = -\tfrac{1}{2}gt^2\mathbf{j} + \mathbf{c}t.$$

The constant vector $\mathbf{c}$ can be determined from a knowledge of the initial launch velocity. Suppose that the launch velocity is $\mathbf{u}$. Substituting $t = 0$ in equation (30) gives $\dot{\mathbf{r}}(0) = \mathbf{u} = \mathbf{c}$, so

$$\mathbf{r}(t) = -\tfrac{1}{2}gt^2\mathbf{j} + \mathbf{u}t.$$

From this equation, we can see that the motion of the projectile is in the vertical plane that contains the launch velocity $\mathbf{u}$, as you might expect. We choose the horizontal x-axis to lie in this plane, as shown in Figures 23 and 24.

To express the launch velocity $\mathbf{u}$ in terms of the unit vectors $\mathbf{i}$ and $\mathbf{j}$, recall from Unit 2 that we can resolve the vector $\mathbf{u}$ into its $\mathbf{i}$-component vector $(|\mathbf{u}|\cos\theta)\,\mathbf{i}$ and its $\mathbf{j}$-component vector $(|\mathbf{u}|\sin\theta)\,\mathbf{j}$ (see Figure 24). Since the launch speed is $|\mathbf{u}| = u$, we have $\mathbf{u} = (u\cos\theta)\,\mathbf{i} + (u\sin\theta)\,\mathbf{j}$. So the solution of equation (29) satisfying these initial conditions is

$$\begin{aligned}\mathbf{r}(t) &= -\tfrac{1}{2}gt^2\mathbf{j} + (u\cos\theta\,\mathbf{i} + u\sin\theta\,\mathbf{j})t \\ &= (ut\cos\theta)\mathbf{i} + (ut\sin\theta - \tfrac{1}{2}gt^2)\mathbf{j}. \qquad (32)\end{aligned}$$

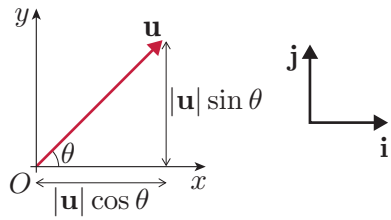


Figure 24 The components of the launch velocity

The vector solution given by equation (32) can be expressed as separate equations for the x- and y-coordinates as

$$x = ut\cos\theta, \qquad (33)$$

$$y = ut\sin\theta - \tfrac{1}{2}gt^2. \qquad (34)$$

So long as the ground is horizontal, we can define the **range** of the projectile to be the horizontal distance between the point of launch and the point of impact. To determine the projectile's range when the launch point and the impact point are in the same horizontal plane, note that the vertical coordinate of its position y will be zero at the launch and again when it hits the ground.

This definition of range is unlikely to be suitable where launch and impact are on an inclined plane, as in the ski-jump example in Exercise 18.

Putting $y = 0$ into equation (34) gives

$$\begin{aligned}0 &= ut\sin\theta - \tfrac{1}{2}gt^2 \\ &= t(u\sin\theta - \tfrac{1}{2}gt).\end{aligned}$$

This equation has two solutions, $t = 0$ and $t = (2u\sin\theta)/g$. At this latter time, the horizontal coordinate of the position x gives the range R of the projectile. From equation (33), this is

$t = 0$ is the instant of launch.

$$R = u\,\frac{2u\sin\theta}{g}\cos\theta = \frac{2u^2\sin\theta\cos\theta}{g}.$$

Now $\sin 2\theta = 2\sin\theta\cos\theta$, so we have

$$R = \frac{u^2\sin 2\theta}{g}. \qquad (35)$$

For a launch on horizontal ground, we must have $\theta > 0$. (For a launch from above ground level, such as from a cliff or bridge, we could have a launch angle below the horizontal, when θ would be negative.)

The sine function never exceeds 1 in value, and we have $\sin 2\theta = 1$ when $2\theta = \frac{\pi}{2}$, that is, when $\theta = \frac{\pi}{4}$. Since the launch angle must be between 0 and $\frac{\pi}{2}$, other solutions can be ignored. So for a given launch speed u, the *maximum range* $R_{\max}$ for a projectile launched from a horizontal surface (ignoring air resistance) is obtained using a launch angle of $\frac{\pi}{4}$ to the horizontal, and this maximum range is

$$R_{\max} = \frac{u^2}{g}.$$

The condition $\dot{y}(t) = 0$ is equivalent to asserting that the vertical component of the velocity is zero when the projectile is at its maximum height.

From equation (34), we see that y is a quadratic function of t, with a negative coefficient of t^2. So the graph of y against t is part of a parabola opening downwards. Such a parabola has a single stationary point where $\dot{y}(t) = 0$, and this will give the maximum value of y. Differentiating equation (34) with respect to t gives

$$\dot{y}(t) = u\sin\theta - gt,$$

and this is 0 when $u\sin\theta - gt = 0$, that is, when $t = (u\sin\theta)/g$. Substituting $t = (u\sin\theta)/g$ into the right-hand side of equation (34), the corresponding *maximum height* is given by

$$\begin{aligned} H &= u\left(\frac{u\sin\theta}{g}\right)\sin\theta - \frac{g}{2}\left(\frac{u\sin\theta}{g}\right)^2 \\ &= \frac{u^2\sin^2\theta}{g} - \frac{u^2\sin^2\theta}{2g} \\ &= \frac{u^2\sin^2\theta}{2g}. \end{aligned} \tag{36}$$

Thus the maximum height occurs when x is half the range R, as you might expect.

Substituting $t = (u\sin\theta)/g$ into the right-hand side of equation (33), the x-component of the projectile's position at the point of maximum height is, using $\sin 2\theta = 2\sin\theta\cos\theta$,

$$\begin{aligned} x &= u\left(\frac{u\sin\theta}{g}\right)\cos\theta \\ &= \frac{u^2\sin\theta\cos\theta}{g} \\ &= \frac{u^2\sin 2\theta}{2g}. \end{aligned}$$

The results derived above apply to any projectile where the points of launch and impact are in the same horizontal plane (and air resistance is ignored). Some problems involving projectiles can conveniently be solved by direct use of these results.

Example 10

During a particular downhill run, a short but sharp rise causes a skier to leave the ground at $25\,\mathrm{m\,s^{-1}}$ at an angle of $\frac{\pi}{6}$ above the horizontal. The ground immediately beyond the rise is horizontal for 60 metres. After this, the slope is again downhill. Will the skier land on the level ground or on the downhill slope beyond it?

Solution

Model the skier as a particle launched from the end of the rise (see Figure 25). With $u = 25$ and $\theta = \frac{\pi}{6}$, the expression in equation (35) for the range on a horizontal surface gives

$$R = \frac{25^2 \sin\frac{\pi}{3}}{9.81} \simeq 55.17.$$

Since this is less than 60 metres, it would seem that the skier will land on the flat part of the run.

(This conclusion is expressed cautiously because of the underlying modelling assumptions. Drag forces may reduce the range of a projectile, but a skier may also experience aerodynamic lift forces that would increase the range. Also, the skier may change the position of her skis relative to the position of her centre of mass – for example, by bending or straightening her legs – which would affect the validity of the model of the skier as a particle.)

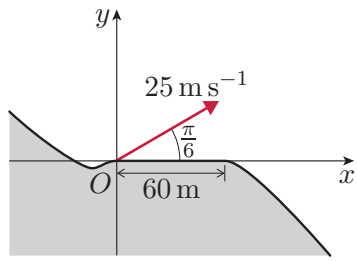


Figure 25 The skier at take-off

This example illustrates how efficiently some questions can be answered by use of general results, such as those giving the range and the maximum height of a projectile when the launch point and the impact point are in the same horizontal plane. When using the results in this subsection in this way, it is important to ensure that they are applicable. We could not, for example, use equation (35) to determine how far the shot putter illustrated in Figure 23(b) would send the shot, since there the point of impact is *not* in the same horizontal plane as that of the launch.

It is important to pay attention to the methodology used in deriving these results, since it can be used in other cases of projectile motion. In the absence of air resistance, we always arrive at the same vector differential equation (29) from Newton's second law. In general, we then need to identify the initial conditions appropriate to the particular situation, find the solution of equation (29) satisfying those initial conditions, and use the solution to address the specific question(s) of interest in the particular problem.

Exercise 25

A ball kicked from a flat piece of ground at an angle of $\frac{\pi}{6}$ above the horizontal lands 40 metres away from where it was kicked. What is the greatest height above the ground that the ball will have reached?

Exercise 26

Consider a projectile launched at an angle θ above the horizontal, from a point at a height h above the origin O, with speed u. Take $t = 0$ to be the moment of launch, and use the coordinate system (and associated unit vectors) shown in Figure 26. Find the solution of $\ddot{\mathbf{r}}(t) = -g\mathbf{j}$ satisfying these initial conditions.

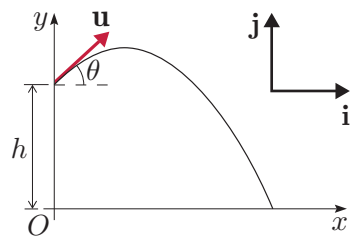


Figure 26 A projectile at launch

Exercise 27

After a road accident, a crashed car is found on a sandy beach at the base of a cliff. The cliff is vertical and is 18 metres high. The investigating police officer finds that the marks in the sand resulting from the car's impact on the beach start 8 metres from the base of the cliff, and that the point of impact is at roughly the same horizontal level as the cliff base. The car appears to have been travelling at right angles to the cliff when it went over. Assuming that the car was travelling in a horizontal direction when it left the cliff, estimate the speed with which it went over.

The vector solution obtained in Exercise 26 can be expressed as separate equations for the components x and y of the position of such a projectile at time t:

$$x = ut\cos\theta, \tag{37}$$

$$y = h + ut\sin\theta - \tfrac{1}{2}gt^2. \tag{38}$$

Equation (37) is identical to equation (33) as the differential equation for x and its initial conditions are unchanged.

As one might expect, launch at a height h simply adds a term h to the y-coordinate.

We now summarise the main results of this subsection.

Projectiles

The equation of motion of a projectile subject only to the force of gravity is

$$\ddot{\mathbf{r}}(t) = -g\mathbf{j}, \tag{39}$$

where $\mathbf{j}$ is a unit vector pointing vertically upwards.

If the projectile is launched at time $t = 0$, from the point $x = 0$, $y = 0$, with launch speed u in the (x, y)-plane and launch angle θ above the horizontal, then the solution of the equation of motion satisfying these initial conditions is

$$x = ut\cos\theta, \tag{40}$$

$$y = ut\sin\theta - \tfrac{1}{2}gt^2. \tag{41}$$

The maximum height H reached by such a projectile is

$$H = \frac{u^2\sin^2\theta}{2g}. \tag{42}$$

If you encounter different initial conditions, you should go back to equation (39) and find the appropriate solution by integration. However, if launch is from $x(0) = 0$, $y(0) = h$, then we need to modify the solutions of the equation of motion (39) given in equations (40) and (41) by adding a term h to the right-hand side of equation (41) as given in equation (38). (In this case, the results for the range and the maximum height are *not* applicable.)

There are also two results for the range of a projectile launched from a level surface.

Range of a projectile launched from a level surface

The range R of a projectile launched from a level surface, subject only to the force of gravity, is

$$R = \frac{u^2 \sin 2\theta}{g}. \tag{43}$$

The maximum range $R_{\max}$ for a launch speed u is achieved with a launch angle $\theta = \frac{\pi}{4}$ and is

$$R_{\max} = \frac{u^2}{g}. \tag{44}$$

Exercise 28

A shot putter launches a shot at a speed of $13\,\mathrm{m\,s^{-1}}$ at an angle of $\frac{\pi}{6}$ above the horizontal from a height of 1.8 metres above ground level. How far will the shot travel in the horizontal direction before it hits the ground, assuming that the ground is horizontal?

Exercise 29

A stone is thrown from a height of 1.5 metres above horizontal ground at an angle of $\frac{\pi}{4}$ above the horizontal and lands at a distance of 30 metres from the point where it was thrown. Estimate the speed with which it was thrown.

5.2 Trajectory of a projectile

A variety of problems can be set about the motion of projectiles. Some can be 'pigeonholed', perhaps requiring you to find the range, or to ensure that a target is hit. Others may require you to bring information about the flight of a projectile to bear on the problem in less predictable ways. We start this subsection with problems that involve hitting a target.

The trajectory of a projectile is the path that it traces. To hit some target, say at P, we require that the point P lies on the trajectory. Suppose that a projectile has launch speed u and launch angle θ above the horizontal, and that it is launched from $(0, 0)$ at time $t = 0$. Then, from work in the previous subsection, we have

A launch angle *below* the horizontal would correspond to a negative value of θ.

$$x = ut\cos\theta, \tag{45}$$

$$y = ut\sin\theta - \tfrac{1}{2}gt^2. \tag{46}$$

If we have no interest in *when* the projectile hits a target, it is efficient to eliminate t to obtain an equation for the trajectory relating y and x directly. Equation (45) gives $t = x/(u\cos\theta)$, and substituting this into equation (46) gives

$$\begin{aligned} y &= u\frac{x}{u\cos\theta}\sin\theta - \frac{g}{2}\left(\frac{x}{u\cos\theta}\right)^2 \\ &= x\tan\theta - x^2\frac{g}{2u^2}\sec^2\theta. \end{aligned}$$

We see that y is a quadratic function of x, and hence that the trajectory is part of a parabola.

As illustrated in Example 11 below, it is often convenient to replace $\sec^2\theta$ by $1+\tan^2\theta$, giving

$$y = x\tan\theta - x^2\frac{g}{2u^2}(1+\tan^2\theta), \tag{47}$$

which is a quadratic equation in $\tan\theta$.

Example 11

A golfer wants to play a recovery shot through a copse of trees. There is a small gap in the foliage at a height of 12 metres and 40 metres in front of him. He knows that with his usual swing, he hits the ball at about $35\,\mathrm{m\,s^{-1}}$. What angle of launch will enable the ball to hit the gap in the foliage?

Solution

We make the usual choice of axes, with origin at the point of launch. The golfer wants the trajectory of the ball to pass through the point $x = 40$, $y = 12$ (working in SI units). So from equation (47) with $u = 35$, we have

$$\begin{aligned} 12 &= 40\tan\theta - 40^2\frac{9.81}{2\times 35^2}(1+\tan^2\theta) \\ &\simeq 40\tan\theta - 6.407(1+\tan^2\theta), \end{aligned}$$

so

$$6.407\tan^2\theta - 40\tan\theta + 18.407 = 0.$$

This is a quadratic equation for $\tan\theta$, with the two solutions

$$\tan\theta = 0.500, \quad \tan\theta = 5.743 \quad \text{to 3 d.p.}$$

Each of these gives a single value for θ in the range $0 \le \theta \le \frac{\pi}{2}$, namely

$$\theta = 0.4638\ (26.58^\circ), \quad \theta = 1.398\ (80.12^\circ).$$

We see that there are two possible launch angles that strike the target (see Figure 27). In this example, the choice of a launch angle of about 27° is perhaps more likely to be suitable, since the other choice would have the ball descending through the foliage at a steep angle, when it would be more likely to hit part of a tree.

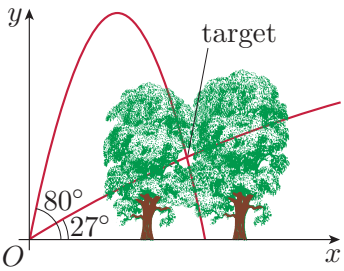


Figure 27 Two launch angles to hit the target

In any problem where we need to find a launch angle (given the launch speed) to hit a specified target, we arrive at a quadratic equation for $\tan\theta$, namely equation (47). So long as this equation has two distinct real roots, there will be two launch angles that enable the target to be hit. Of course, the roots may coincide. In either case, we say that the target is **achievable**. On the other hand, we may arrive at a quadratic equation with no real roots. In this case, the target is not achievable – for the given launch speed, there is no launch angle that enables the target to be hit.

Exercise 30

Suppose that a projectile has launch speed u at an angle θ above the horizontal, and that it is launched from $(0, h)$ at time $t = 0$. (That is, the projectile is launched at a height h above the origin.) By eliminating t from equations (37) and (38), obtain an equation (relating y to x) for the trajectory of such a projectile.

We see from Exercise 30 that for a launch at height h above the origin, we need to add only a term h to the right-hand side of equation (47) for the trajectory. This result is frequently useful.

Trajectory of a projectile

For a projectile launched at time $t = 0$ from $(0, h)$ with launch speed u and launch angle θ above the horizontal, the trajectory has the equation

$$y = h + x\tan\theta - x^2\frac{g}{2u^2}\sec^2\theta. \tag{48}$$

It is often convenient to use the trigonometric identity $\sec^2\theta = 1 + \tan^2\theta$ to give

$$y = h + x\tan\theta - x^2\frac{g}{2u^2}(1 + \tan^2\theta), \tag{49}$$

which is a quadratic equation in $\tan\theta$.

Exercise 31

A basketball player is 2.6 metres (horizontally) from the basket. The basket is 3 metres above ground level. The player launches the ball at $7\,\mathrm{m\,s^{-1}}$, from a height of 1.8 metres above ground level. What angle of launch should the player choose?

We now look at an example that requires more work to answer the posed question.

Example 12

This example concerns baseball fielders throwing the ball back to the catcher. Assume throughout that the point of launch and the point of impact are at the same horizontal level.

(a) A fielder can just throw a ball a distance of 60 metres. How fast can the fielder throw the ball?

(b) A fielder needs to throw a ball to the catcher from a distance of 58 metres. Assuming that the fielder throws directly to the catcher at the speed calculated in part (a), what is the shortest time in which the fielder can return the ball to the catcher?

(c) Suppose that a second fielder is midway between the first fielder and the catcher (so that each gap is 29 metres), and that each fielder throws at the speed calculated in part (a). The first fielder throws to the second fielder, then the second fielder throws to the catcher. As well as the time in flight, the second fielder requires 0.3 seconds to catch and throw the ball. Would this 'relaying' result in a quicker return of the ball to the catcher?

Solution

(a) In the previous subsection we found that the maximum range of a projectile (for launch speed u) is achieved at a launch angle $\frac{\pi}{4}$, and this maximum range is u^2/g (see equation (44)). So for the fielder, we have $u^2/g = 60$, which gives

$$u = \sqrt{60g} = \sqrt{60 \times 9.81}$$
$$= 24.26 \quad \text{to 2 d.p.}$$

So the fielder can throw the ball at a speed of approximately $24.3\,\mathrm{m\,s^{-1}}$.

(b) Taking the point of launch as origin, and working in metres, the trajectory of the ball needs to pass through the point $(58, 0)$. So using equation (47), we need a launch angle θ where

$$0 = 58\tan\theta - 58^2\frac{g}{2 \times 60g}(1 + \tan^2\theta).$$

It is simpler, as well as more accurate, to use $u = \sqrt{60g}$ here.

This can be rearranged as

$$\tan^2\theta - \frac{120}{58}\tan\theta + 1 = 0.$$

This quadratic equation has solutions $\tan\theta = 0.7696$ and $\tan\theta = 1.299$. The corresponding values of θ (between 0 and $\frac{\pi}{2}$) are 0.656 (37.6°) and 0.915 (52.4°). Throwing the ball at either of these angles will return it to the catcher.

To find the time that it takes the ball to reach the catcher, we can use equation (45). When $x = 58$, the time t must satisfy

$$58 = ut\cos\theta = \sqrt{60g}\,t\cos\theta.$$

We obtain different times depending on the choice of the launch angle θ. With $\theta = 0.656$, the time is 3.02 seconds. With $\theta = 0.915$, the time is 3.92 seconds. As one might expect, the lower angle of launch gives the shorter flight time, so the fastest possible return time is 3.02 seconds.

(c) The total time to return the ball to the catcher is $2T + 0.3$ seconds, where T is the time (in seconds) to throw the ball 29 metres at the launch speed calculated in part (a). To ensure that the ball passes through the point $(29, 0)$, we need a launch angle θ satisfying equation (47) with these coordinates, that is,

$$0 = 29\tan\theta - 29^2\frac{g}{2\times 60g}(1+\tan^2\theta),$$

so

$$\tan^2\theta - \frac{120}{29}\tan\theta + 1 = 0.$$

This has solutions $\tan\theta = 0.2577$ and $\tan\theta = 3.880$. The corresponding launch angles are 0.252 (14.5°) and 1.319 (75.6°). Flight times are 1.23 seconds and 4.79 seconds. The lower launch angle gives the shorter flight time, and the total return time to the catcher for the 'relay' is

$$2\times 1.23 + 0.3 = 2.77.$$

This time of 2.77 seconds *is* shorter than the time found in part (b) for a direct throw.

In the previous subsection we saw that for a launch speed u on a horizontal surface, the maximum range for a projectile is obtained with a launch angle $\frac{\pi}{4}$ and is u^2/g (equation (44)). We next consider the angle of launch required to give the maximum range when launch is from *above* ground level.

We continue to define the range to be the horizontal displacement between the point of launch and the point of impact, even for a launch above ground level. The following calculation of the maximum range brings together ideas about projectiles and methods from calculus. The approaches that we use are chosen to minimise the complexity of the algebra; they are not always those that first come to mind.

This horizontal displacement measures the length of a shot put, for example.

Consider a projectile launched with speed u from a point at height h above the ground, which we assume is horizontal. We choose the origin and axes as shown in Figure 28, and $t = 0$ as the time of launch. If the launch angle is θ above the horizontal, the equation of the trajectory is given by equation (49) as

$$y = h + x\tan\theta - x^2\frac{g}{2u^2}(1+\tan^2\theta).$$

Let R be the range. Then since $(R, 0)$ lies on the trajectory, we have

$$0 = h + R\tan\theta - R^2\frac{g}{2u^2}(1+\tan^2\theta).$$

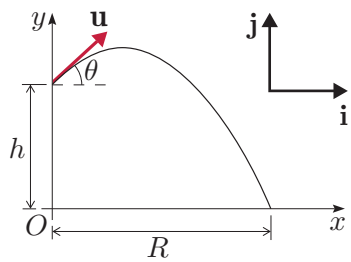


Figure 28 The range of the projectile

Defining $z = \tan\theta$ and $L = u^2/g$ gives

$$0 = h + Rz - \frac{R^2}{2L}(1 + z^2). \tag{50}$$

We assume that $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. Remember that a negative value for θ corresponds to a launch angle below the horizontal.

For a given launch speed and height, h and L are constants. We want to maximise the range R by choice of the launch angle θ or, equivalently, by choice of z. The global maximum of a function can occur at a stationary point or at an endpoint of its domain. However, the endpoints of this domain, $\theta = -\frac{\pi}{2}$ and $\theta = \frac{\pi}{2}$, both lead to a range of $R = 0$ (the motion is straight up and down), so the maximum range must occur at a stationary point. In order to simplify the algebra, we consider R as a function of z, so we want to find values of z for which $dR/dz = 0$.

Now implicit differentiation of equation (50) with respect to z gives

$$0 = \frac{d}{dz}\big(z\,R(z)\big) - \frac{1}{2L}\,\frac{d}{dz}\big((1 + z^2)(R(z))^2\big)\,,$$

or

$$0 = \left(R + z\,\frac{dR}{dz}\right) - \frac{1}{2L}\left(2zR^2 + (1 + z^2)2R\,\frac{dR}{dz}\right).$$

Setting $dR/dz = 0$ reduces this equation to

$$0 = R - \frac{zR^2}{L}.$$

So the maximum range occurs when $z = L/R$. Substituting $z = L/R$ into equation (50) gives

$$\begin{aligned} 0 &= h + R\,\frac{L}{R} - \frac{R^2}{2L}\left(1 + \frac{L^2}{R^2}\right) \\ &= h + L - \frac{R^2}{2L} - \frac{L}{2} \\ &= h + \frac{L}{2} - \frac{R^2}{2L}. \end{aligned}$$

Hence

$$R = \sqrt{L^2 + 2Lh}.$$

This maximum range is achieved when

$$\tan\theta = z = \frac{L}{R} = \frac{L}{\sqrt{L^2 + 2Lh}} = \frac{1}{\sqrt{1 + 2h/L}},$$

that is, when

$$\theta = \arctan\left(\frac{1}{\sqrt{1 + 2h/L}}\right).$$

Strictly speaking, we have not shown that the range given above *is* a maximum. All we have shown is that it is a stationary value, which could also be a minimum value, for example. However, physically we know that

the projectile does have a maximum range, and as $z = L/R$ is the only stationary point, this stationary point must be a maximum. Alternatively, we could show mathematically that the stationary point given by $z = L/R$ is a maximum by considering the sign of the second derivative d^2R/dz^2.

Maximum range for an elevated launch

A projectile launched with speed u from a height h above ground level has a maximum range on horizontal ground given by

$$R_{\max} = \sqrt{L^2 + 2Lh}, \tag{51}$$

where $L = u^2/g$. This maximum range is achieved using the launch angle

$$\theta = \arctan\left(\frac{1}{\sqrt{1 + 2h/L}}\right). \tag{52}$$

If $h = 0$, equation (51) gives $R_{\max} = L$, as it should, since L is the maximum range for a launch from ground level (when $h = 0$). Also, if $h > 0$, then $R_{\max} > L$. As one might expect, a launch from above ground level achieves a maximum range greater (for the same launch velocity) than a launch at ground level. More generally, the greater the height of the launch point, the greater the maximum range.

With $h = 0$, equation (52) gives $\theta = \arctan 1 = \frac{\pi}{4}$, again corresponding to our previous result for a launch from ground level.

Exercise 32

At a tutorial, one of the students, who happens to be an expert shot putter, asserts that aiming to launch at an angle $\frac{\pi}{4}$ has always been good enough for him. He says that improving launch speed is the key to good shot putting. Assume that the student launches the shot from a height of 2 metres above ground level.

(a) The student can put a shot 17 metres with a launch angle $\frac{\pi}{4}$ above the horizontal. Calculate the speed at which the shot is being launched to achieve this range.

(b) For launch at the speed calculated in part (a), use equations (51) and (52) to find the optimum launch angle and the corresponding range.

(c) If the student achieves a launch speed 1% higher than that calculated in part (a) and launches at an angle $\frac{\pi}{4}$ above the horizontal, what range will he achieve?

As usual in this unit, model the ball as a particle. This means that any swerve that a footballer may achieve by kicking the ball with spin will be overlooked.

Exercise 33

A footballer taking a free kick launches the ball from ground level so that it just clears a player who is 10 metres away and 2 metres high. The ball enters the goal 30 metres away at a height of 2.4 metres.

(a) Take as the origin the point from which the ball was kicked. Let the launch speed be u, and let the launch angle above the horizontal be θ. Use equation (48) (with $h = 0$) twice to obtain two equations that u and θ must satisfy.

Multiply one of these equations by a suitable constant, so that the term $(g \sec^2 \theta)/2u^2$ has the same coefficient in each equation. Then subtract one equation from the other to eliminate u, and thus obtain an equation that is satisfied by $\tan \theta$. Hence find the launch angle θ.

(b) At what speed was the ball kicked? What period of time elapsed from the moment the ball was kicked until it entered the goal?

Learning outcomes

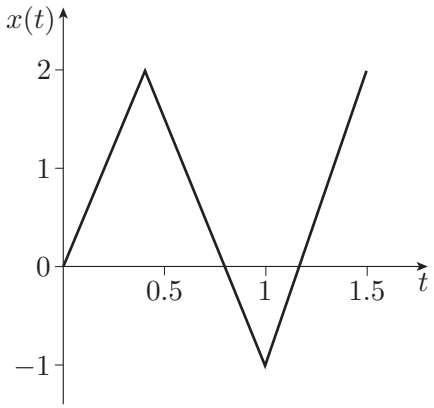
After studying this unit, you should be able to:

- understand and use the basic terms for the description of the motion of particles: position, velocity and acceleration
- understand and use vector functions
- understand the fundamental laws of Newtonian mechanics
- solve mechanics problems in one dimension by drawing a sketch, choosing a suitable x-axis and origin, making assumptions, drawing a force diagram, applying Newton's second law, taking the x-component, and making suitable substitutions
- solve mechanics problems in one dimension that involve one or more of the forces of gravity, friction and air resistance
- understand the concept of terminal speed, and use it in solving mechanics problems in one dimension
- apply Newton's second law in vector form to problems in more than one dimension
- solve problems relating to the motion of a projectile in the absence of air resistance.

Solutions to exercises

Solution to Exercise 1

(a) The graph of x against t for the puck is shown below.



(b) (i) The distance between the first player and the back wall is the distance between where the puck starts and where it first changes direction. This can be read off the graph as 2 m.

(ii) The second player must be further away than the first player since the x-coordinate is -1 when the puck changes direction for a second time, indicating that the second player is 3 m from the wall.

(iii) The speed at which the puck is travelling is given by the slope of the distance–time graph. The slope after the second player hits the puck is greater than the slope after the first player hits the puck, so the second player gives the puck more speed.

Solution to Exercise 2

(a) $\dfrac{d\mathbf{r}(t)}{dt} = \left(\dfrac{d}{dt}t^2\right)\mathbf{i} + \left(\dfrac{d}{dt}10t\right)\mathbf{j} = 2t\mathbf{i} + 10\mathbf{j}.$

So the velocity of the particle at $t = 1$ is

$$\mathbf{v} = 2\mathbf{i} + 10\mathbf{j},$$

and the speed is

$$|\mathbf{v}| = \sqrt{4 + 100} = \sqrt{104} \simeq 10.20 \quad \text{to 2 d.p.}$$

(b) $\dfrac{d\mathbf{r}(t)}{dt} = (\cos t)\mathbf{i} - (\sin t)\mathbf{j} + \mathbf{k}.$

So the velocity of the particle at $t = 1$ is

$$\mathbf{v} = (\cos 1)\mathbf{i} - (\sin 1)\mathbf{j} + \mathbf{k},$$

and the speed is

$$|\mathbf{v}| = \sqrt{\cos^2 1 + \sin^2 1 + 1} = \sqrt{2} \simeq 1.41 \quad \text{to 2 d.p.}$$

Solution to Exercise 3

(a) The graph for particle A is its position–time graph. The component of velocity of the particle along the x-axis, dx/dt, is given by the slope of this graph: it can be seen that the slope starts at a high value and decreases as time increases. So particle A is travelling fastest when $t = 0$.

(b) The graph for particle B shows the variation with time of the particle's component of velocity along the x-axis. For the particle to change direction, this velocity must change from positive to negative or from negative to positive (i.e. the graph must cross the t-axis). The graph shown does not do this, so particle B does not change its direction.

(c) The graph for particle C shows the variation with time of the particle's component of acceleration along the x-axis. On the interval $[0, 2.5]$ the acceleration is positive, so the velocity increases in this interval. On the interval $[2.5, 5]$ the acceleration is negative, so the velocity decreases in this interval. To discover whether the particle is ever stationary for $t > 0$, we need to know whether its velocity along the x-axis is ever zero for $t > 0$. Now, since acceleration is obtained by differentiating velocity, we can obtain velocity by integrating acceleration. We also know that the definite integral of a function over a given interval gives the 'area' under the graph of the function, where 'areas' below the t-axis are negative. So the 'area' under the graph gives us the velocity.

Now, on the interval $[0, 2.5]$, the 'area' under the curve is positive, and since the particle starts from rest, the velocity after 2.5 seconds is positive. However, on the interval $[2.5, 5]$, the 'area' under the curve is negative, and furthermore the magnitude of this 'area' is greater than the magnitude of the 'area' for the interval $[0, 2.5]$; therefore the velocity after 5 seconds is negative. Hence, since the particle has both positive and negative velocity in the given time interval, it must be momentarily stationary at some point towards the end of the time interval.

Solution to Exercise 4

Since

$$a(t) = \frac{dv}{dt} = 18t - 20,$$

we have

$$v = \int (18t - 20)\, dt = 9t^2 - 20t + A,$$

where A is a constant.

Using the initial condition $v(0) = 3$, we obtain $A = 3$. Hence the component of the velocity of the particle along the x-axis is given by

$$v(t) = 9t^2 - 20t + 3.$$

Now $v(t) = dx/dt$, so

$$x = \int (9t^2 - 20t + 3)\, dt = 3t^3 - 10t^2 + 3t + B,$$

where B is a constant.

The initial condition $x(0) = 7$ gives $B = 7$. Hence the component of the position of the particle along the x-axis is given by

$$x(t) = 3t^3 - 10t^2 + 3t + 7.$$

Substituting $t = 10$ into the expressions for $x(t)$ and $v(t)$ gives

$$x(10) = 3000 - 1000 + 30 + 7 = 2037,$$
$$v(10) = 9 \times 10^2 - 20 \times 10 + 3 = 703,$$

so at time $t = 10$ the particle is 2037 metres from the origin, with a speed of $703\,\mathrm{m\,s^{-1}}$ along the positive x-axis. Thus the position is $\mathbf{r}(10) = 2037\,\mathbf{i}$ and the velocity is $\mathbf{v}(10) = 703\,\mathbf{i}$.

Solution to Exercise 5

We have

$$a(t) = \frac{dv}{dt} = ge^{-kt}.$$

Integrating this gives

$$v = \int ge^{-kt}\, dt = A - \frac{g}{k}e^{-kt},$$

where A is a constant.

The initial condition $v(0) = 0$ gives $A = g/k$, so the velocity is given by

$$v(t) = \frac{g}{k} - \frac{g}{k}e^{-kt} = \frac{g}{k}(1 - e^{-kt}).$$

Then from $v(t) = dx/dt$ we have

$$x = \int \left(\frac{g}{k} - \frac{g}{k}e^{-kt}\right) dt = \frac{g}{k}t + \frac{g}{k^2}e^{-kt} + B,$$

where B is a constant.

The initial condition $x(0) = 0$ gives $B = -g/k^2$, so the position is given by

$$x(t) = \frac{g}{k}t + \frac{g}{k^2}e^{-kt} - \frac{g}{k^2} = \frac{g}{k}t - \frac{g}{k^2}(1 - e^{-kt}).$$

Therefore the velocity and position of the particle are given by the vector functions

$$\mathbf{v}(t) = \frac{g}{k}\left(1 - e^{-kt}\right)\mathbf{i},$$
$$\mathbf{r}(t) = \left(\frac{g}{k}t - \frac{g}{k^2}(1 - e^{-kt})\right)\mathbf{i}.$$

Solution to Exercise 6

Given that $a(t) = a_0$, the relationship $a = v\,dv/dx$ gives

$$v\frac{dv}{dx} = a_0.$$

Applying the separation of variables method to this differential equation gives

$$\int v\,dv = \int a_0\,dx,$$

so

$$\tfrac{1}{2}v^2 = a_0x + C,$$

where C is a constant.

Using the initial condition that the velocity is v_0 along the x-axis at $x = x_0$ gives $C = \frac{1}{2}v_0^2 - a_0x_0$, so

$$\tfrac{1}{2}v^2 = a_0x + \tfrac{1}{2}v_0^2 - a_0x_0.$$

Multiplying through by 2 and rearranging gives

$$v^2 = v_0^2 + 2a_0(x - x_0),$$

as required.

Solution to Exercise 7

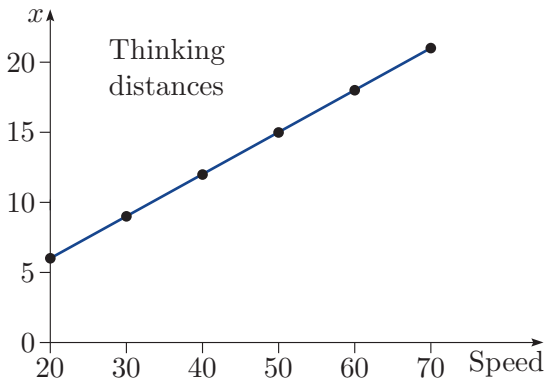
(a) To find v in terms of t, substitute dv/dt for a; then v may be found by direct integration. To find x in terms of t, substitute dx/dt for v and integrate again.

(b) To find a relationship between v and x, substitute $v\,dv/dx$ for a. The result is an equation that can be solved by separation of variables, whose solution will give the required relationship.

(c) To find x in terms of t, substitute d^2x/dt^2 for a, and dx/dt for v. The result is a linear constant-coefficient second-order differential equation for the variable x, namely $\ddot{x} + 3\dot{x} + 2x = \cos t$. This can be solved by the methods of Unit 1. (The general solution is $x = Ae^{-2t} + Be^{-t} + \frac{1}{10}\cos t + \frac{3}{10}\sin t$.)

Solution to Exercise 8

(a) (i) The figure below shows the thinking distance in metres against the speed in miles per hour. The speed of a car before the thinking phase is the value given in the table. The speed after the thinking phase is exactly the same, because the driver has not yet reacted to the hazard. So the acceleration is zero during this phase, and with $x_0 = 0$, the formula $x = x_0 + v_0t + \frac{1}{2}a_0t^2$ reduces to $x = v_0t$.



The values of v_0 and the thinking distance x, in SI units, can be calculated from the values given in the table (using the conversion factor given in the margin next to the question); the only unknown is the thinking time t. The value of t for each pair of speeds and distances can be calculated from $t = x/v_0$.

The calculation, given that the thinking distance at 20 mph is 6 m, and using the given conversion factor, is (in seconds)

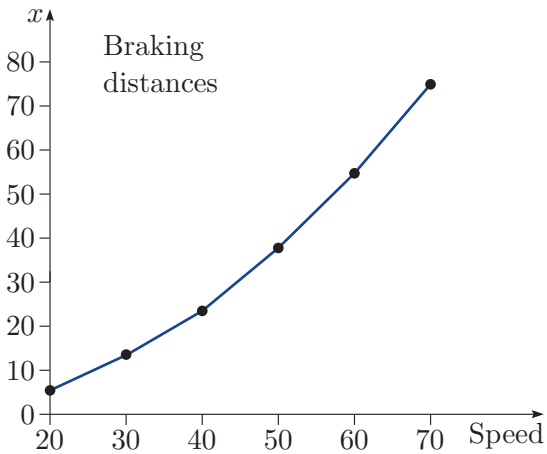
$$t = \frac{x}{v_0} = \frac{6}{20 \times 0.447} \simeq 0.67.$$

(Other pairs give the same value for the thinking time.)

So the model used for calculating the thinking distance data is

$$x = v_0 t = 0.67 v_0.$$

(ii) The figure below shows the braking distance in metres against the speed in miles per hour. The relationship is clearly not linear. The speed of a car at the start of the braking phase is the speed at the end of the thinking phase (i.e. the value given in the table, converted to SI units). The speed at the end of the braking phase is zero ($v = 0$). Assuming that the braking is uniform, so that the acceleration is constant, we can use the formula $v^2 = v_0^2 + 2a_0x$, which for $v = 0$ reduces to $0 = v_0^2 + 2a_0x$.



The values of v_0 and the braking distance x, in SI units, can be calculated from the values given in the table; the only unknown is the acceleration a_0. The value of a_0 for each pair of speeds and distances can be calculated from $a_0 = -v_0^2/(2x)$. (Note that from this equation the acceleration is negative, which is a good check because the car is stopping!)

The calculation for the last pair of values given in the table is

$$a_0 = -\frac{v_0^2}{2x} = -\frac{(70 \times 0.447)^2}{2 \times 75} \simeq -6.53.$$

(Other pairs give values between $-6.42\,\mathrm{m\,s^{-2}}$ and $-6.66\,\mathrm{m\,s^{-2}}$ for the acceleration, which is about two-thirds of the magnitude of the acceleration due to gravity.)

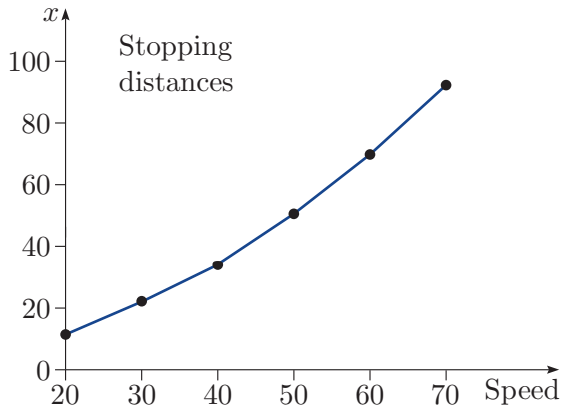
So the model used for calculating the braking distance data is

$$x \simeq -\frac{v_0^2}{2a_0} = \frac{v_0^2}{13.1}.$$

(b) The overall stopping distance is equal to the thinking distance plus the braking distance, and these distances are calculated separately using the models in part (a) to give

$$x = 0.67v_0 + \frac{v_0^2}{13.1}.$$

The figure below shows the stopping distance in metres against the speed in miles per hour.



At 45 mph, the total stopping distance is calculated as

$$\begin{aligned} x &= 0.67 \times (45 \times 0.447) + \frac{(45 \times 0.447)^2}{13.1} \\ &\simeq 13.5 + 30.9 \\ &= 44.4 \quad \text{to 1 d.p.} \end{aligned}$$

Thus the stopping distance at 45 mph is approximately 44.4 m. (Note that the calculated distance is nearly halfway between the tabulated values for 40 mph and 50 mph – it is not *exactly* halfway, because the braking distance is a quadratic function of v_0.)

Solution to Exercise 9

Any moving car is subject to resistive forces, namely air resistance, the internal frictional forces in the car's engine, transmission and wheel bearings, and the external frictional forces between the car's tyres and the road. In order to maintain a constant velocity, it is necessary to apply a motive force that balances these resistive forces.

Solution to Exercise 10

The component of the force of gravity in the direction of a unit vector pointing down the slope is non-zero. If the slope is steep enough, then this force down the slope will be greater than the resistive force of friction, causing the toboggan to accelerate.

Solution to Exercise 11

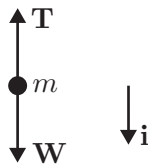
Newton's second law, $\mathbf{F} = m\mathbf{a}$, is a vector equation. Taking the magnitude of both sides of the equation (and using the fact that mass is always positive) gives the scalar equation $|\mathbf{F}| = m|\mathbf{a}|$, into which the values given in the question can be substituted (after converting the mass from grams into the SI unit kilograms) to obtain

$$10 = 0.2|\mathbf{a}|.$$

This gives $|\mathbf{a}| = 50$, so the force produces an acceleration of magnitude $50\,\mathrm{m\,s^{-2}}$.

Solution to Exercise 12

(a) The only forces on the hanging object are the weight $\mathbf{W}$ of the object (a downward force, in the direction of the positive x-axis shown) and the tension force $\mathbf{T}$ due to the string (an upward force, in the direction of the negative x-axis). The resultant force on the object is the sum $\mathbf{W} + \mathbf{T}$ of these forces. This information is shown in the force diagram in the margin.



(b) Applying Newton's second law to this system gives

$$m\mathbf{a} = \mathbf{W} + \mathbf{T}.$$

From the force diagram, we have $\mathbf{W} = |\mathbf{W}|\,\mathbf{i} = mg\mathbf{i}$, where g is the magnitude of the acceleration due to gravity. Similarly, the tension force due to the string is $\mathbf{T} = -|\mathbf{T}|\,\mathbf{i}$. The given acceleration of the hanging object is $1\,\mathrm{m\,s^{-2}}$ downwards, so $\mathbf{a} = \mathbf{i}$. With this information we can resolve the equation above in the $\mathbf{i}$-direction to obtain

$$m \times 1 = mg - |\mathbf{T}|.$$

Substituting $m = 10$ and rearranging gives

$$|\mathbf{T}| = 10g - 10 \simeq 88.1 \quad (\text{using } g = 9.81\,\mathrm{m\,s^{-2}}).$$

So $\mathbf{T} = -|\mathbf{T}|\,\mathbf{i} = -88.1\mathbf{i}$, that is, the tension force due to the string is 88.1 newtons in the upward direction.

Solution to Exercise 13

Substitute the given force and mass into the equation for Newton's second law to obtain

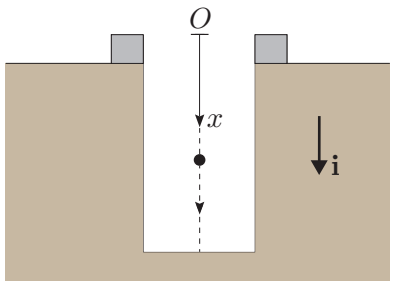
$$4000\,\mathbf{a} = 50\,000\,\mathbf{i} + 60\,000\,\mathbf{j} + 100\,000\,\mathbf{k}.$$

So $\mathbf{a} = 12.5\mathbf{i} + 15\mathbf{j} + 25\mathbf{k}$, which has magnitude (in $\mathrm{m\,s^{-2}}$) $\sqrt{12.5^2 + 15^2 + 25^2} \simeq 32$. This is well below the threshold of $6g \simeq 59\,\mathrm{m\,s^{-2}}$, so the pilot should remain conscious.

Solution to Exercise 14

◀ Draw picture ▶

(a) First, we draw a picture.



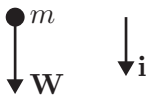
◀ Choose axes ▶

The x-axis is chosen to point vertically downwards, with the origin O at the top of the well, as shown above.

◀ State assumptions ▶

The stone is modelled as a particle, and the model assumes that the only force is the stone's weight due to gravity. The force diagram is as follows.

◀ Draw force diagram ▶



◀ Apply Newton's 2nd law ▶

Applying Newton's second law to the stone gives $\mathbf{W} = m\mathbf{a}$. Since $\mathbf{W} = mg\mathbf{i}$, we have $m\mathbf{a} = mg\mathbf{i}$, and resolving in the $\mathbf{i}$-direction gives

$$a = g.$$

◀ Solve differential equation ▶

Using $a = dv/dt$, we obtain

$$\frac{dv}{dt} = g.$$

Integrating this gives

$$v = gt + A,$$

where A is a constant.

The initial condition that the stone is dropped from rest ($v = 0$ when $t = 0$) gives $A = 0$. Hence

$$v = gt.$$

Now using $v = dx/dt$, we have

$$\frac{dx}{dt} = gt.$$

Integrating this gives

$$x = \tfrac{1}{2}gt^2 + B,$$

where B is a constant.

The initial condition $x = 0$ when $t = 0$ gives $B = 0$. So

$$x = \tfrac{1}{2}gt^2.$$

Using this equation with $t = 3$ gives ◀ Interpret solution ▶

$$x = \tfrac{1}{2} \times 9.81 \times 3^2 = 44.145.$$

So the well is estimated to be about 44 metres deep.

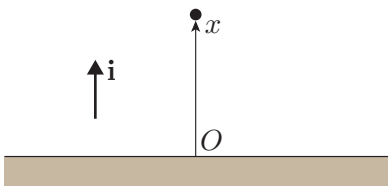
(b) Using the equation for v with $t = 3$ gives

$$v = gt = 9.81 \times 3 = 29.43.$$

So the predicted speed of the stone as it reaches the bottom is about $29\,\mathrm{m\,s^{-1}}$.

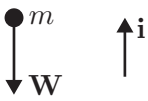
Solution to Exercise 15

(a) First, we draw a picture. ◀ Draw picture ▶



Choose the x-axis to point vertically upwards, with its origin O at the point from which the ball is thrown, as shown above. ◀ Choose axes ▶

The model assumes that the weight due to gravity is the only force acting on the ball, so the force diagram is as follows. ◀ State assumptions ▶ ◀ Draw force diagram ▶



Applying Newton's second law to the ball gives ◀ Apply Newton's 2nd law ▶

$$m\mathbf{a} = \mathbf{W}.$$

Since the x-axis points upwards, the weight of the ball is given by $\mathbf{W} = -mg\mathbf{i}$. The acceleration is downwards, so $\mathbf{a} = a\mathbf{i}$, where a is negative. Resolving in the $\mathbf{i}$-direction gives $ma = -mg$. Dividing by the mass gives

$$a = -g.$$

◀ Solve differential equation ▶

To answer the question, we need an equation relating x to t, and another relating v to t or to x. Since the acceleration $a = -g$ is constant, one approach is to use the constant acceleration formulas of Subsection 1.2. The initial velocity is $10\,\mathrm{m\,s^{-1}}$ upwards from the origin. Using the notation of Subsection 1.2, the initial velocity is $\mathbf{v}_0 = 10\mathbf{i}$, so $v_0 = 10$ and $a_0 = -g$, and hence

$$v = v_0 + a_0 t = 10 - gt,$$
$$x = v_0 t + \tfrac{1}{2} a_0 t^2 = 10t - \tfrac{1}{2} g t^2.$$

(The approach of Example 4, using the substitutions $a = dv/dt$ and $v = dx/dt$ and integrating, leads to the same pair of equations.)

◀ Interpret solution ▶

The ball reaches its maximum height when $v = 0$, and from the equation for v this occurs at time

$$t = \frac{10}{g} \simeq 1.02 \quad \text{to 2 d.p.}$$

Since the motion started at $t = 0$, the duration of the upward flight of the ball is about 1 second.

(b) Substituting $t = 10/g$ into the equation for x, we have

$$x = \frac{100}{g} - \frac{100}{2g} \simeq 5.10 \quad \text{to 2 d.p.}$$

So the maximum height attained by the ball is about 5.1 metres.

The motion up and down has the same acceleration $-g\mathbf{i}$, so it can be treated as one motion. There is no need to consider the upward and downward motions separately.

(c) The ball reaches the ground when $x = 0$, and from the equation for x this occurs when

$$0 = 10t - \tfrac{1}{2} g t^2 = t(10 - \tfrac{1}{2} g t).$$

Hence $t = 0$ or $t = 20/g \simeq 2.04$ to two decimal places. Now $t = 0$ corresponds to the time when the ball is thrown, so the ball *returns* to the ground after approximately 2 seconds.

(d) Substituting $t = 20/g$ into the equation for v gives

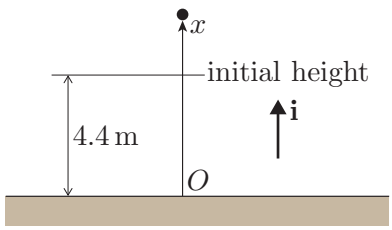
$$v = 10 - gt = 10 - 20 = -10, \quad \text{so } \mathbf{v} = 10(-\mathbf{i}).$$

The ball reaches the ground on its return with a speed of $10\,\mathrm{m\,s^{-1}}$. (The $(-\mathbf{i})$ indicates that the ball is now travelling in the direction of decreasing x, i.e. downwards.)

Solution to Exercise 16

◀ Draw picture ▶

(a) First, we draw a picture.



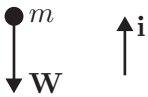
The x-axis is chosen to point vertically upwards, with the origin O at ground level, as shown in the diagram. (The other obvious choice for the origin, which you may well have chosen, is at the point where the ball is thrown. Both choices of course lead to the same answers below.)

◀ Choose axes ▶

The model assumes that the only force acting on the ball is its weight due to gravity, so the force diagram is as follows.

◀ State assumptions ▶

◀ Draw force diagram ▶



Applying Newton's second law to the ball gives $m\mathbf{a} = \mathbf{W}$, where $\mathbf{a} = a\mathbf{i}$ and $\mathbf{W} = -mg\mathbf{i}$ (as the x-axis points upwards). Resolving in the $\mathbf{i}$-direction gives

◀ Apply Newton's 2nd law ▶

$$ma = -mg, \quad \text{so } a = -g.$$

We want equations relating x to t, and v to x or t. One approach is to use the substitution $a = dv/dt$, to obtain

◀ Solve differential equation ▶

$$\frac{dv}{dt} = -g.$$

Integrating this, we obtain

$$v = -gt + A,$$

where A is a constant.

The initial velocity of the ball is $7.6\mathbf{i}$, so $v = 7.6$ when $t = 0$, which gives $A = 7.6$. So

$$v = -gt + 7.6.$$

Therefore using the substitution $v = dx/dt$ gives

$$\frac{dx}{dt} = -gt + 7.6.$$

Integrating this gives

$$x = -\tfrac{1}{2}gt^2 + 7.6t + B,$$

where B is a constant.

The initial condition that $x = 4.4$ when $t = 0$ leads to $B = 4.4$. So

$$x = -\tfrac{1}{2}gt^2 + 7.6t + 4.4.$$

(Since the acceleration is constant, another approach is to use the constant acceleration formulas to obtain the equations for v and x. Note that with the choice of origin used here, $x_0 = 4.4$, $v_0 = 7.6$ and $a_0 = -9.81$. You could use the equation $x = x_0 + v_0t + \frac{1}{2}a_0t^2$ to obtain the equation for x.)

The ball reaches the ground when $x = 0$. Substituting this into the equation for x gives a quadratic equation for the time t,

◀ Interpret solution ▶

$$4.905t^2 - 7.6t - 4.4 = 0,$$

whose solution is

$$t = \frac{7.6 \pm \sqrt{(-7.6)^2 - 4 \times 4.905 \times (-4.4)}}{2 \times 4.905},$$

so $t \simeq 1.998$ or $t \simeq -0.4489$. The negative time is before the ball is thrown and may therefore be ignored.

So the ball lands about 2 seconds after being thrown.

It may be easier to find the speed v at ground level by solving

$v\,dv/dx = -g$

to obtain

$\frac{1}{2}v^2 - \frac{1}{2}v_0^2 = -g(x - x_0)$,

so when $x = 0$,

$v^2 = v_0^2 + 2gx_0$.

(b) Substituting this time into the equation for v gives

$$v = -gt + 7.6 = -9.81 \times 1.998 + 7.6 \simeq -12.00 \quad \text{to 2 d.p.}$$

So the ball lands with a speed of about $12.0\,\mathrm{m\,s^{-1}}$. (The negative sign for v confirms that the ball is moving downwards, i.e. the speed is $12.0\,\mathrm{m\,s^{-1}}$ in the $-\mathbf{i}$ direction.)

Solution to Exercise 17

(a) Choose the same x-axis and the same origin as in Example 4, and proceed in exactly the same way as before until you reach the equation $a = g$. Now write a as $v\,dv/dx$ to obtain

$$v\frac{dv}{dx} = g.$$

Solving this differential equation by the method of separation of variables, we have

$$\int v\,dv = \int g\,dx, \quad \text{so} \quad \tfrac{1}{2}v^2 = gx + A,$$

where A is a constant.

Now the marble starts from rest, so $v = 0$ when $x = 0$, which leads to $A = 0$. Hence $\frac{1}{2}v^2 = gx$, or equivalently,

$$v = \sqrt{2gx},$$

where we have taken the positive square root because the velocity is positive throughout the motion.

(This equation could also have been obtained from the constant acceleration formula (8).)

(b) Putting $v = dx/dt$ in the equation for v gives

$$\frac{dx}{dt} = \sqrt{2gx}.$$

Again we use the method of separation of variables to solve this differential equation. So we have

$$\int 1\,dt = \int \frac{1}{\sqrt{2gx}}\,dx = \frac{1}{\sqrt{2g}}\int x^{-1/2}dx = \frac{1}{\sqrt{2g}}\,2x^{1/2} + B,$$

where B is a constant. So

$$t = \sqrt{\frac{2x}{g}} + B.$$

The marble starts at the origin, so $x = 0$ when $t = 0$, which gives $B = 0$. Hence

$$t = \sqrt{\frac{2x}{g}}.$$

(This equation could also have been obtained from the constant acceleration formula (7).)

(c) At $x = 77$, the equation for t yields

$$t = \sqrt{\frac{2x}{g}} = \sqrt{\frac{2 \times 77}{9.81}} \simeq 3.962,$$

and the equation for v gives

$$v = \sqrt{2gx} = \sqrt{2 \times 9.81 \times 77} \simeq 38.87.$$

We take the positive square root as the motion is in the **i**-direction.

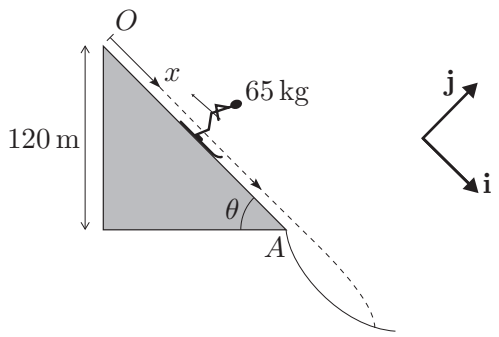
So the object hits the water after about 3.96 seconds, with a speed of about $38.9\,\mathrm{m\,s^{-1}}$. (The answers are, of course, the same as those in Example 4.)

Solution to Exercise 18

(a) As the question asks for the speed for two different angles of the slope of the jump, it is sensible (as suggested in Figure 15) to use θ to be an arbitrary angle of slope and substitute for θ at the interpretation stage.

A diagram of the situation is shown with the question as Figure 15, and is repeated here.

◀ Draw picture ▶

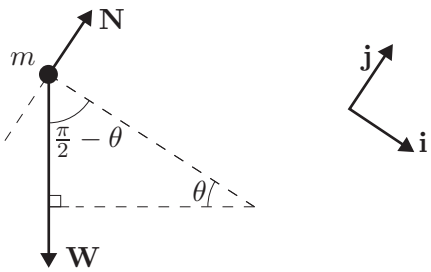


The x-axis is given as pointing down the slope with origin O at the top of the slope, as shown above. The model assumes that the skier can be treated as a particle, and that the only forces acting on the skier are her weight due to gravity and the normal reaction from the slope.

◀ Choose axes ▶

◀ State assumptions ▶

◀ Draw force diagram ▶



◀ Apply Newton's 2nd law ▶

Applying Newton's second law to the skier gives

$$m\mathbf{a} = \mathbf{W} + \mathbf{N}.$$

From the force diagram, $\mathbf{N} = |\mathbf{N}|\,\mathbf{j}$. We can resolve $\mathbf{W}$ into components:

$$\begin{aligned}\mathbf{W} &= |\mathbf{W}|\cos(\tfrac{\pi}{2} - \theta)\,\mathbf{i} - |\mathbf{W}|\sin(\tfrac{\pi}{2} - \theta)\,\mathbf{j}\\ &= mg\sin\theta\,\mathbf{i} - mg\cos\theta\,\mathbf{j}.\end{aligned}$$

Now we can resolve in the **i**-direction to obtain

$$ma = mg\sin\theta, \quad \text{so} \quad a = g\sin\theta.$$

◀ Solve differential equation ▶

As we want the velocity when the skier has travelled a vertical distance of 120 metres, it is best to find v as a function of x. Since the acceleration is constant, we can use equation (8) to obtain

$$v^2 = v_0^2 + 2ax = 0 + 2(g\sin\theta)x,$$

so

$$v = \sqrt{2gx\sin\theta}.$$

◀ Interpret solution ▶

The task is to find the velocity when the skier has travelled a vertical distance of 120 metres. Now x is the distance travelled down the slope; so, using trigonometry, $x = 120/\sin\theta$. Substituting for x in the equation for v gives

$$v = \sqrt{2g \times \frac{120}{\sin\theta} \times \sin\theta} = \sqrt{240g} \simeq 48.52 \quad \text{to 2 d.p.}$$

So the speed of the skier at the bottom of the slope angled at $\frac{\pi}{4}$ is about $48.5\,\mathrm{m\,s^{-1}}$.

The substitution of θ at an early stage would have led to a lot of extra work here as v is independent of θ.

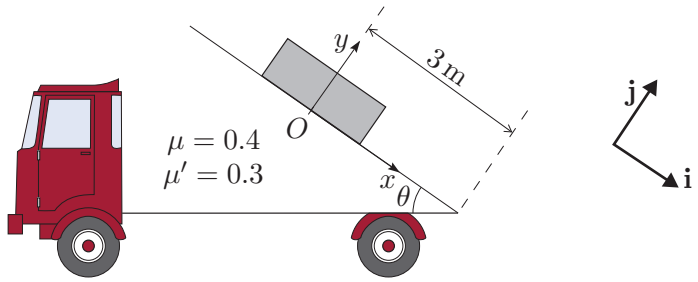
(b) In part (a), all mention of θ cancelled from the final expression. So the answer remains the same: the speed at the bottom of the $\frac{\pi}{3}$ slope is about $48.5\,\mathrm{m\,s^{-1}}$.

(In the equation for v, note that $x\sin\theta = h$ is the vertical height thus the speed at the bottom depends on the height of the slope and not the angle of the slope.)

Solution to Exercise 19

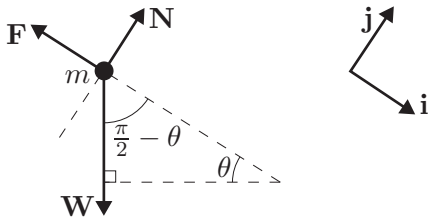
◀ Draw picture ▶

(a) Let θ be the angle that the carrier makes with the horizontal, as shown below, and let m be the mass of the block.



Choose the axes to be parallel and perpendicular to the carrier, with origin O at the rest position of the concrete block, as shown above. ◀ Choose axes ▶

The block is modelled as a particle, and we assume that the only forces acting on the block are the block's weight $\mathbf{W}$, the normal reaction with the carrier $\mathbf{N}$, and the friction force up the slope $\mathbf{F}$. This gives the following force diagram. ◀ State assumptions ▶ ◀ Draw force diagram ▶



The equilibrium condition for the block is ◀ Apply law(s) ▶

$$\mathbf{F} + \mathbf{N} + \mathbf{W} = \mathbf{0}.$$

When the block is on the point of moving, the magnitude of the friction force is given by

$$|\mathbf{F}| = \mu|\mathbf{N}|.$$

From the force diagram, $\mathbf{N} = |\mathbf{N}|\,\mathbf{j}$ and $\mathbf{F} = -|\mathbf{F}|\,\mathbf{i}$. The weight can be resolved into components as ◀ Solve equation(s) ▶

$$\begin{aligned}\mathbf{W} &= |\mathbf{W}|\cos\left(\tfrac{\pi}{2} - \theta\right)\,\mathbf{i} + |\mathbf{W}|\sin\left(\tfrac{\pi}{2} - \theta\right)\,(-\mathbf{j}) \\ &= mg\sin\theta\,\mathbf{i} - mg\cos\theta\,\mathbf{j}.\end{aligned}$$

Now we can resolve the equilibrium equation in the **i**-direction, giving

$$-|\mathbf{F}| + 0 + mg\sin\theta = 0,$$

so

$$|\mathbf{F}| = mg\sin\theta.$$

Resolving in the **j**-direction gives

$$0 + |\mathbf{N}| - mg\cos\theta = 0,$$

so

$$|\mathbf{N}| = mg\cos\theta.$$

But when the block is on the point of slipping, the friction equation applies, and substituting the values of $|\mathbf{F}|$ and $|\mathbf{N}|$ gives

$$mg\sin\theta = \mu mg\cos\theta.$$

Rearranging this gives

$$\tan\theta = \mu.$$

Substituting $\mu = 0.4$ into this equation gives $\theta = \arctan 0.4 \simeq 0.381$ to three decimal places. ◀ Interpret solution ▶

So the angle at which the concrete block begins to slide is about 0.38 radians or about 22°.

(b) To solve the dynamics problem when the block is in motion down the carrier, we start in exactly the same way as for the statics problem. So we start the analysis of the motion by applying Newton's second law, using the same axes, with the origin O at the point at which slipping first occurs.

◀ Apply Newton's 2nd law ▶

Applying Newton's second law to the block gives

$$m\mathbf{a} = \mathbf{F} + \mathbf{N} + \mathbf{W}.$$

The acceleration is down the carrier, so $\mathbf{a} = a\mathbf{i}$ and all the forces are resolved in exactly the same way as in the statics problem. So we can resolve in the $\mathbf{i}$-direction to obtain

$$ma = mg\sin\theta - |\mathbf{F}|.$$

Now from the moment at which the block begins to slide, $|\mathbf{F}| = \mu'|\mathbf{N}|$, so this equation becomes

$$ma = mg\sin\theta - \mu'|\mathbf{N}|.$$

To find $|\mathbf{N}|$, we resolve the original equation in the $\mathbf{j}$-direction to obtain

$$0 = -mg\cos\theta + |\mathbf{N}|.$$

Therefore $|\mathbf{N}| = mg\cos\theta$ (as before). Thus

$$ma = mg\sin\theta - \mu' mg\cos\theta,$$

so

$$a = g\sin\theta - \mu' g\cos\theta.$$

◀ Solve differential equation ▶

For a fixed angle θ, the acceleration is constant, so we can use the constant acceleration formulas from Section 1. Since we want to relate time to distance travelled, the appropriate formula is equation (7):

$$x = x_0 + v_0 t + \tfrac{1}{2}a_0 t^2.$$

Initially, the block is at rest at $x_0 = 0$, so $v_0 = 0$ and this becomes

$$x = \tfrac{1}{2}g(\sin\theta - \mu'\cos\theta)t^2.$$

◀ Interpret solution ▶

The time before the block slides off the back of the truck is calculated from this equation using the value of θ calculated in part (a). Substituting the distance travelled ($x = 3$) and $\mu' = 0.3$, we obtain

$$3 \simeq \tfrac{1}{2}g[\sin(0.381) - 0.3\cos(0.381)]t^2 \simeq 0.458t^2,$$

Using the exact angle $\theta = \arctan 0.4$ gives $t = 2.57$ to two decimal places.

so $t \simeq 2.56$ to two decimal places.

So once the concrete block begins to slide, it takes about 2.6 seconds to slide off the back of the truck.

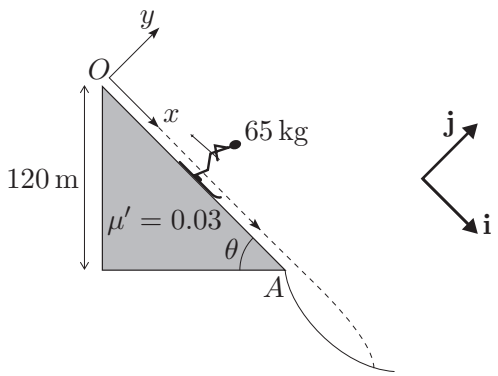
Solution to Exercise 20

◀ Draw picture ▶

First, we draw a picture.

◀ Choose axes ▶

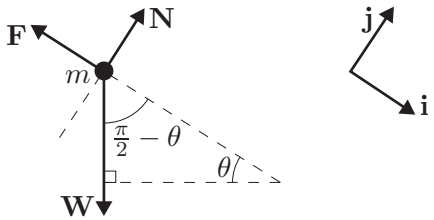
The x-axis is given to be parallel to the slope, with origin O at the top of the slope. Choose the y-axis to be perpendicular to the slope, as shown below.



The skier is modelled as a particle, and we assume that the only forces acting on the skier are her weight, the normal reaction with the slope, and the friction force up the slope. This gives the following force diagram.

◀ State assumptions ▶

◀ Draw force diagram ▶



Applying Newton's second law to the skier gives

◀ Apply Newton's 2nd law ▶

$$m\mathbf{a} = \mathbf{W} + \mathbf{N} + \mathbf{F}.$$

From the force diagram, $\mathbf{N} = |\mathbf{N}|\,\mathbf{j}$ and $\mathbf{F} = -|\mathbf{F}|\,\mathbf{i}$. The weight can be resolved into components as

$$\begin{aligned}\mathbf{W} &= |\mathbf{W}|\cos(\tfrac{\pi}{2} - \theta)\,\mathbf{i} + |\mathbf{W}|\sin(\tfrac{\pi}{2} - \theta)\,(-\mathbf{j}) \\ &= mg\sin\theta\,\mathbf{i} - mg\cos\theta\,\mathbf{j}.\end{aligned}$$

The acceleration is down the slope, so $\mathbf{a} = a\mathbf{i}$, and we can resolve in the $\mathbf{i}$-direction to obtain

$$ma = mg\sin\theta + 0 - |\mathbf{F}|.$$

Using $|\mathbf{F}| = \mu'|\mathbf{N}| = 0.03|\mathbf{N}|$, this becomes

$$ma = mg\sin\theta - 0.03|\mathbf{N}|.$$

Resolving in the $\mathbf{j}$-direction leads to

$$0 = -mg\cos\theta + |\mathbf{N}| + 0,$$

so $|\mathbf{N}| = mg\cos\theta$. Substituting gives

$$ma = mg\sin\theta - 0.03 \times mg\cos\theta,$$

so

$$a = g\sin\theta - 0.03g\cos\theta.$$

Using equation (8), we obtain

◀ Solve differential equation ▶

$$v^2 = v_0^2 + 2ax = 0 + 2(g\sin\theta - 0.03g\cos\theta)x,$$

so

$$v = \sqrt{2g(\sin\theta - 0.03\cos\theta)x}.$$

◀ Interpret solution ▶

As in Exercise 18, $x = 120/\sin\theta$, so

$$v = \sqrt{2g(\sin\theta - 0.03\cos\theta)\frac{120}{\sin\theta}}$$
$$= \sqrt{240g - 7.2g\cot\theta}.$$

Substituting $\theta = \frac{\pi}{4}$ into this equation gives $v \simeq 47.79$ to two decimal places.

So the speed of the skier at the bottom of the slope angled at $\frac{\pi}{4}$ is about $47.8\,\mathrm{m\,s^{-1}}$.

Substituting $\theta = \frac{\pi}{3}$ into the equation gives $v \simeq 48.10$ to two decimal places.

Note that the calculations are easier here since θ was used in developing and solving the model. The solution can be used to answer the question for the two different values of θ.

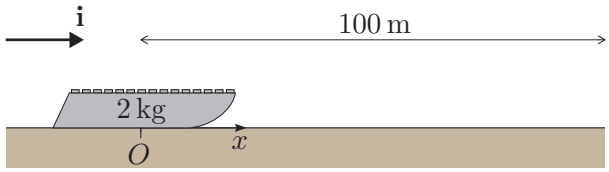
So the speed of the skier at the bottom of the slope angled at $\frac{\pi}{3}$ is about $48.1\,\mathrm{m\,s^{-1}}$.

From these answers, it can be seen that the steeper the slope, the faster the final speed of the skier. The final speed is always less than the $48.5\,\mathrm{m\,s^{-1}}$ calculated in Exercise 18, which omitted friction from the model.

Solution to Exercise 21

◀ Draw picture ▶

First, we draw a picture.



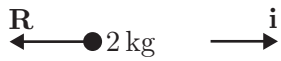
◀ Choose axes ▶

Choose the x-axis along the direction of motion, with origin O at the point where the toboggan has speed $2\,\mathrm{m\,s^{-1}}$, as shown in the diagram.

◀ State assumptions ▶

We model the toboggan as a particle. There are three forces acting on the toboggan: its weight, the normal reaction from the horizontal surface, and the air resistance force. The vertical forces **W** and **N** are in equilibrium and do not affect the motion. This leaves only the air resistance force **R**, which opposes the motion, as shown in the force diagram below.

◀ Draw force diagram ▶



◀ Apply Newton's 2nd law ▶

Applying Newton's second law gives

$$m\mathbf{a} = \mathbf{R} + \mathbf{N} + \mathbf{W} = \mathbf{R}.$$

The question states that the quadratic air resistance model should be used, so $\mathbf{R} = -c_2D^2|\mathbf{v}|\mathbf{v} = -c_2D^2v^2\mathbf{i}$, since $\mathbf{v} = v\mathbf{i}$, $v > 0$, thus $|\mathbf{v}| = v$. Resolving in the **i**-direction gives

$$ma = -c_2D^2v^2.$$

◀ Solve differential equation ▶

The question requires the velocity of the toboggan after 100 m, so we use the substitution $a = v\,dv/dx$ to obtain

$$mv\frac{dv}{dx} = -c_2 D^2 v^2.$$

Solving by separation of variables,

$$\int \frac{1}{v}\,dv = -\int k\,dx,$$

where $k = c_2 D^2/m$, so

$$\ln v = -kx + A,$$

where A is a constant. Hence

$$v = Be^{-kx},$$

where $B = e^A$ is another constant. The initial condition that the toboggan is initially moving at $2\,\mathrm{m\,s^{-1}}$ (i.e. $v = 2$ when $x = 0$) gives $B = 2$, so

$$v = 2e^{-kx},$$

where $k = c_2 D^2/m$.

Substitute the data given in the question to find the velocity after 100 m: ◀ Interpret solution ▶

$$k = \frac{c_2 D^2}{m} = \frac{0.2 \times (0.05)^2}{2} = 2.5 \times 10^{-4},$$
$$v(100) = 2\exp(-2.5 \times 10^{-4} \times 100) \simeq 1.95.$$

So the percentage decrease in velocity is about

$$100\,\frac{v(0) - v(100)}{v(0)} = 100\,\frac{2 - 1.95}{2} = 2.5,$$

that is, after 100 metres the toboggan has lost only about 2.5% of its speed.

Solution to Exercise 22

First suppose that the linear air resistance model applies, so that the terminal speed is

$$v_\mathrm{T} = \frac{mg}{c_1 D} = \frac{65 \times 9.81}{1.7 \times 10^{-4} \times 10} \simeq 4 \times 10^5.$$

So $Dv_\mathrm{T} \simeq 4 \times 10^6$, which is greater than 10^{-5}, so the linear model does not apply.

Now suppose that the quadratic model applies, so that the terminal speed is

$$v_\mathrm{T} = \sqrt{\frac{mg}{c_2 D^2}} = \sqrt{\frac{65 \times 9.81}{0.2 \times 10^2}} \simeq 5.6.$$

So $Dv_\mathrm{T} \simeq 56$, which is greater than 1, so the quadratic model does not apply either.

The condition for the quadratic model is closer to being satisfied than the condition for the linear model, so the quadratic model is likely to produce the better estimate. So the conclusion is that the landing speed of the parachutist is approximately $6\,\mathrm{m\,s^{-1}}$.

(Looking at Figure 19, it can be seen that $6\,\mathrm{m\,s^{-1}}$ (with $D = 10$, so that $\log_{10}(D|\mathbf{v}|) \simeq 1.8$) falls to the right of the range of validity of the quadratic model. In this region, you can see that the quadratic model lies just above the experimental curve, so that it gives a slight overestimate of the air resistance. So it should be expected that the actual landing speed is slightly greater than $6\,\mathrm{m\,s^{-1}}$. Also note that this landing speed would be achieved by a particle falling under gravity, without air resistance, from a height of 1.8 m.)

Solution to Exercise 23

You saw in Exercise 22 that the quadratic air resistance model is better than the linear model for problems of this type. Rearrangement of the quadratic model $v_{\mathrm{T}} = \sqrt{mg/c_2D^2}$ gives

$$D = \sqrt{\frac{mg}{c_2 v_{\mathrm{T}}^2}} = \sqrt{\frac{70 \times 9.81}{0.2 \times 13^2}} \simeq 4.5.$$

So the effective diameter needs to be at least 4.5 metres.

As in the solution to Exercise 22, we have $\log_{10}(D|\mathbf{v}|) = \log_{10}(4.5 \times 13) \simeq 1.8$, so the quadratic model overestimates the air resistance in this case. So to be safe, a parachute with effective diameter of 5 or even 6 metres would probably be needed.

Solution to Exercise 24

(a) The equation of motion of the marble is given in the question as

$$v\frac{dv}{dx} = g - kv^2,$$

where $k = c_2D^2/m$ is a positive constant.

Proceeding as in Example 8, we solve the differential equation by separation of variables:

$$\int \frac{v}{g - kv^2}\,dv = \int 1\,dx.$$

Now the first difference due to the changing initial conditions occurs, since if $v = 50$, then

$$g - kv^2 = 9.81 - \frac{0.2 \times (0.02)^2 \times (50)^2}{0.013} \simeq -5.6.$$

So the denominator of the first integrand is negative (whereas it was positive before). Rewriting this to make the denominator positive by taking out the factor -1 gives

$$-\int \frac{v}{kv^2 - g}\,dv = \int 1\,dx,$$

so

$$-\frac{1}{2k}\ln(kv^2 - g) = x + A,$$

where A is a constant.

To determine A, use the initial condition that $v = v_0$ when $x = 0$:

$$-\frac{1}{2k}\ln(kv_0^2 - g) = 0 + A.$$

Substituting for A and rearranging gives

$$-2kx = \ln(kv^2 - g) - \ln(kv_0^2 - g),$$

so

$$e^{-2kx} = \frac{kv^2 - g}{kv_0^2 - g}$$

or

$$e^{-2kx}(kv_0^2 - g) = kv^2 - g,$$

thus

$$kv^2 = g + e^{-2kx}(kv_0^2 - g).$$

So

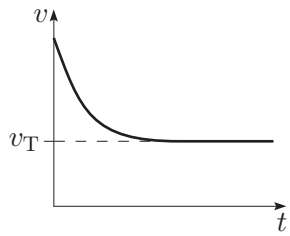
$$v = \sqrt{\frac{g}{k} + e^{-2kx}\frac{kv_0^2 - g}{k}}.$$

(b) The eventual behaviour of the velocity $\mathbf{v} = v\mathbf{i}$ can be seen from the equation for kv^2 above, since as x becomes large, the exponential term will become vanishingly small. So as x becomes large, v satisfies the equation

$$g = kv^2.$$

This means that the speed will tend to $\sqrt{g/k}$ as x becomes large.

This is the same eventual behaviour as in Example 8, that is, the marble's speed tends to a terminal speed. Moreover, the terminal speed $\sqrt{g/k}$ has the same value as for the marble falling from rest. Now, however, the speed *decreases* exponentially towards the terminal speed. The behaviour is shown in the sketch in the margin.



Solution to Exercise 25

In the text, we obtained equation (35) for the range, namely $R = u^2 \sin 2\theta/g$. With $\theta = \frac{\pi}{6}$, the kick has range 40 metres, so the launch speed u must satisfy

$$u^2 = \frac{40g}{\sin\frac{\pi}{3}} = \frac{80g}{\sqrt{3}}.$$

We also obtained the expression (36) for the greatest height $H = (u^2 \sin^2\theta)/2g$ reached by a projectile. With $\theta = \frac{\pi}{6}$ and $u^2 = 80g/\sqrt{3}$, this gives

$$H = \frac{80g \times \sin^2\frac{\pi}{6}}{2g\sqrt{3}} = \frac{10}{\sqrt{3}} \simeq 5.77 \quad \text{to 2 d.p.}$$

So the greatest height reached by the ball is about 5.8 metres.

Solution to Exercise 26

We need to find the solution of $\ddot{\mathbf{r}}(t) = -g\mathbf{j}$ satisfying $\dot{\mathbf{r}}(0) = \mathbf{u}$, where $\mathbf{u} = (u\cos\theta)\mathbf{i} + (u\sin\theta)\mathbf{j}$ and $\mathbf{r}(0) = h\mathbf{j}$ (since the point of launch is $(0, h)$). The integral of $\ddot{\mathbf{r}}(t) = -g\mathbf{j}$ is

$$\dot{\mathbf{r}}(t) = -gt\mathbf{j} + \mathbf{c},$$

where $\mathbf{c}$ is a constant vector.

Substituting $t = 0$ and using the initial condition $\dot{\mathbf{r}}(0) = \mathbf{u}$, we must have $\mathbf{c} = \mathbf{u}$.

Integrating again gives

$$\mathbf{r}(t) = -\tfrac{1}{2}gt^2\mathbf{j} + \mathbf{u}t + \mathbf{d},$$

where $\mathbf{d}$ is a constant vector.

Substituting $t = 0$ and using the initial condition $\mathbf{r}(0) = h\mathbf{j}$, we must have $\mathbf{d} = h\mathbf{j}$.

Hence the required solution is

$$\begin{aligned}\mathbf{r}(t) &= -\tfrac{1}{2}gt^2\mathbf{j} + (ut\cos\theta)\mathbf{i} + (ut\sin\theta)\mathbf{j} + h\mathbf{j}\\ &= (ut\cos\theta)\mathbf{i} + (h + ut\sin\theta - \tfrac{1}{2}gt^2)\mathbf{j}.\end{aligned}$$

Solution to Exercise 27

We choose the y-axis vertically upwards and the x-axis horizontal. We choose the origin at the bottom of the cliff at beach level. Suppose that the car left the cliff with a speed u, travelling in a horizontal direction (at right angles to the cliff). Then the initial conditions are $\mathbf{r}(0) = 18\mathbf{j}$ and $\dot{\mathbf{r}}(0) = u\mathbf{i}$. We can use the solution of $\ddot{\mathbf{r}}(t) = -g\mathbf{j}$ derived in Exercise 26, with $\theta = 0$, which is

$$\mathbf{r}(t) = ut\mathbf{i} + (h - \tfrac{1}{2}gt^2)\mathbf{j},$$

or, separated into components,

$$x = ut,$$
$$y = h - \tfrac{1}{2}gt^2.$$

We know that $x = 8$ when $y = 0$ (assuming that the car hit the ground exactly 8 metres from the cliff and modelling the car as a particle). So from the first equation, the car hit the ground at $t = 8/u$. Substituting this into the second equation gives

$$0 = 18 - \tfrac{1}{2}g(64/u^2).$$

Thus, as u is positive,

$$u = \sqrt{\frac{64 \times 9.81}{2 \times 18}} \simeq 4.18 \quad \text{to 2 d.p.}$$

So the car left the cliff at a speed of just over $4\,\mathrm{m\,s^{-1}}$.

Solution to Exercise 28

Taking the origin to be at ground level, and using equations (37) and (38), the position of the shot at a time t after the launch is given by

$$x = 13t\cos\tfrac{\pi}{6},$$
$$y = 1.8 + 13t\sin\tfrac{\pi}{6} - \tfrac{1}{2}\times 9.81t^2.$$

To find the time when the shot hits the ground, we substitute $y = 0$ in the second equation and solve the resulting quadratic equation for t. The solutions are $t = 1.560$ and $t = -0.2352$. The negative solution represents a time before the shot is launched and so can be rejected. At $t = 1.560$, we have $x = 17.57$.

So the shot lands at a horizontal distance of 17.57 metres from the point of launch.

Solution to Exercise 29

Taking the origin to be at ground level, we can use equations (37) and (38) with $h = 1.5$ and $\theta = \frac{\pi}{4}$. Suppose that the launch speed is u. Then equation (37) gives $x = ut\cos\frac{\pi}{4} = ut/\sqrt{2}$. If the stone hits the ground when $t = T$, we have $30 = uT/\sqrt{2}$, so $T = 30\sqrt{2}/u$. We know that $y = 0$ when $t = T$, so substituting into equation (38) gives

$$\begin{aligned} 0 &= 1.5 + u\,\frac{30\sqrt{2}}{u}\,\frac{1}{\sqrt{2}} - \frac{g}{2}\left(\frac{30\sqrt{2}}{u}\right)^2 \\ &= 31.5 - \frac{900g}{u^2}. \end{aligned}$$

This gives $u = 30\sqrt{9.81/31.5} \simeq 16.74$.

So the launch speed was approximately $16.74\,\mathrm{m\,s^{-1}}$.

Solution to Exercise 30

Equations (37) and (38) are

$$x = ut\cos\theta,$$
$$y = h + ut\sin\theta - \tfrac{1}{2}gt^2.$$

From the first equation, $t = x/(u\cos\theta)$. Substituting this into the second equation gives

$$\begin{aligned} y &= h + u\,\frac{x}{u\cos\theta}\sin\theta - \frac{g}{2}\left(\frac{x}{u\cos\theta}\right)^2 \\ &= h + x\tan\theta - x^2\frac{g}{2u^2}\sec^2\theta. \end{aligned}$$

Alternatively, using $\sec^2\theta = 1 + \tan^2\theta$,

$$y = h + x\tan\theta - x^2\frac{g}{2u^2}(1 + \tan^2\theta).$$

Solution to Exercise 31

We choose the origin to be at ground level, vertically below the point of launch. So the equation of the trajectory of the basketball is equation (49) with $h = 1.8$ and $u = 7$ (using SI units). In order for the ball to pass through the hoop, we want the point $x = 2.6$, $y = 3$ to be on this trajectory. Hence

$$3 = 1.8 + 2.6\tan\theta - (2.6)^2\frac{9.81}{2\times 7^2}(1+\tan^2\theta).$$

This simplifies to the quadratic equation

$$0.6767\tan^2\theta - 2.6\tan\theta + 1.877 = 0.$$

(Alternatively, and more efficiently, you may have chosen the origin to be the point from which the ball was launched. However, this leads to the same equation for $\tan\theta$.)

This equation for $\tan\theta$ has the two solutions

$$\tan\theta = 2.879 \quad \text{and} \quad \tan\theta = 0.963.$$

Each of these gives a single value for θ in the range $0 \leq \theta \leq \frac{\pi}{2}$:

$$\theta = 1.236\ (70.9^\circ) \quad \text{and} \quad \theta = 0.767\ (43.9^\circ).$$

We see that there are two possible launch angles that enable the target to be hit. In this example, the choice of a launch angle of approximately 71° is more likely to be suitable, since this has the ball descending towards the net at the steeper angle, so the ball is less likely to catch on the rim of the basket.

Solution to Exercise 32

(a) We choose the origin to be at ground level, vertically below the point of launch. So the equation of the trajectory of the shot is equation (48) with $h = 2$ and $\theta = \frac{\pi}{4}$ (using SI units). The trajectory must pass through the point of impact, namely $x = 17$, $y = 0$. So the launch speed u must satisfy the equation

$$0 = 2 + 17 - \frac{17^2\times 9.81\times 2}{2u^2}.$$

(Alternatively, you may have chosen the origin to be the point from which the shot is launched. Then the equation of the trajectory is equation (48) with $h = 0$, and the point of impact is $x = 17$, $y = -2$. However, you should arrive at the same equation for u as above.)

We have $u^2 = (17^2 \times 9.81)/19$, so $u \simeq 12.22$.

So the launch speed is about $12.22\,\mathrm{m\,s^{-1}}$.

(b) With u as calculated in part (a), the parameter L in equations (51) and (52) has the value $u^2/g \simeq 15.21$. We also have $h = 2$, so the value of θ giving the maximum range is (from equation (52))

$$\begin{aligned}\theta &= \arctan\left(\frac{1}{\sqrt{1+2h/L}}\right)\\ &= \arctan\left(\frac{1}{\sqrt{1+4/15.21}}\right)\\ &= 0.727 \ (41.7^\circ).\end{aligned}$$

The range achieved with this optimum launch angle is (from equation (51))

$$\begin{aligned}R &= \sqrt{L^2+2Lh}\\ &= \sqrt{15.21^2+4\times 15.21}\\ &\simeq 17.09.\end{aligned}$$

So the optimum launch angle is about 41.7°, with a range of approximately 17.1 metres.

(c) An improvement of 1% on the launch speed $12.22\,\mathrm{m\,s^{-1}}$ calculated in part (a) would give launch speed $12.22 \times 1.01 = 12.34\,\mathrm{m\,s^{-1}}$. With this launch speed and launch angle $\frac{\pi}{4}$, choosing the origin to be at ground level, vertically below the point of launch, the equation of the trajectory of the shot (equation (48)) is

$$y = 2 + x - x^2\frac{9.81}{2\times 12.34^2}\times 2.$$

At the point of impact $y = 0$, which leads to the quadratic equation

$$0 = 2 + x - 0.0644x^2.$$

This has solutions $x = 17.32$ and $x = -1.79$.

We can reject the negative solution, which represents the point behind the putter where the trajectory intersects ground level. So the range of the put will be approximately 17.3 metres.

(We can see from the answers to parts (b) and (c) that the student is right in saying that a small increase in launch speed is more effective in increasing the range than is getting the optimum launch angle. However, although slight, the improvement in range (of 9 cm) resulting from putting at the optimum angle could be the difference between winning and coming nowhere! So one might as well try to achieve the optimum launch angle.)

Solution to Exercise 33

(a) The trajectory must pass through the points $(10, 2)$ and $(30, 2.4)$. So using equation (48) twice, we have

$$2 = 10\tan\theta - 100\,\frac{g}{2u^2}\sec^2\theta,$$

$$2.4 = 30\tan\theta - 900\,\frac{g}{2u^2}\sec^2\theta.$$

To eliminate the $\sec^2\theta$ term, we multiply the first equation by 9 to obtain

$$18 = 90\tan\theta - 900\,\frac{g}{2u^2}\sec^2\theta.$$

Subtracting gives

$$15.6 = 60\tan\theta.$$

So $\tan\theta = 0.26$ and $\theta = 0.2544$ $(14.6°)$.

(b) Substituting $\theta = 0.2544$ into the first equation in part (a) gives

$$2 = 10\tan(0.2544) - \frac{50g}{u^2}\sec^2(0.2544),$$

so

$$u^2 = \frac{50 \times 9.81\sec^2(0.2544)}{10\tan(0.2544) - 2} = 872.28$$

and $u \simeq 29.53$.

So the ball was kicked at a speed of approximately $29.5\,\mathrm{m\,s^{-1}}$ to one decimal place.

Using equation (45), we have $x = ut\cos\theta$. We know that the ball entered the goal when $x = 30$, so $t = 30/(29.53\cos(0.2544)) \simeq 1.05$ to two decimal places.

So just over 1 second after having been kicked, the ball entered the goal.

Index

absolute error 29, 30
acceleration 184, 189, 193
acceleration due to gravity 125
accumulation 4
achievable target 223
addition of vectors 101, 103
    in component form 108
air resistance 205, 206
algebraic rules for scaling and adding vectors 103
analytic solution 3
angle between vectors 115
approximate solution 24
arbitrary constant 7
area
    of a parallelogram 122
    of a triangle 122
associated homogeneous equation 46
associativity 103
auxiliary equation 37

balanced rod 142
birth rate 4
boundary condition 62
boundary value 62
boundary-value problem 62

Cartesian components of a vector 108
Cartesian coordinates 100, 105
Cartesian unit vectors 100, 107
centre of mass 141, 182
characteristic equation 37
choosing axes 125
coefficient of dynamical friction 202
coefficient of kinetic friction 202
coefficient of sliding friction 202, 204
coefficient of static friction 130, 132
commutativity 102, 103, 112
complementary function 46
component form
    of a cross product 121
    of a dot product 114
    of a vector 107, 108
component of a vector 108
    in an arbitrary direction 116
constant acceleration 189
constant-coefficient equation 34
cost 31
cross product 118, 120

death rate 4
decay constant 15
dependent variable 6

derivative 6
    of a vector function 183
differentiable function 7
differential equation 7
    first-order 7
    general solution 8
    homogeneous 16
    inhomogeneous 16
    linear 16, 34
    non-homogeneous 16
    order 7
    particular solution 8
    solution 7
direct integration 11
direction field 21
direction of a vector 97, 99
discriminant 44
displacement 95
distributivity 103, 112, 120
domain 10
dot product 111, 112
dynamics 96, 177

effective diameter 211
efficiency 31
equal vectors 99
equation of a straight line 110
equation of motion 187
equilibrium 124, 144
equilibrium condition
    for a particle 124
    for a rigid body 144
Euler, Leonhard 27
Euler's method 27
exceptional cases in the method of undetermined coefficients 57
exponential function 51
extended body 141

first-order differential equation 7
force 95, 193, 194
    mathematical representation 96
force diagram 124
force of gravity 124
friction 202
friction force 129
fundamental laws of Newtonian mechanics 194

general solution of a differential equation 8, 43, 46, 47
gradient 21
gravitational force 124

hinge 146

homogeneous differential equation 16, 34

inclined plane 132
indefinite integral 11
independent variable 6
inhomogeneous differential equation 16, 34, 44
initial condition 9, 60
initial value 9, 60
initial-value problem 9, 60
input 4
input–output principle 4
integrating factor 19
integrating factor method 19

kinematics 179

launch angle 216
launch velocity 215
law of addition of forces 194
law of addition of mass 194
laws of motion 95
Leibniz, Gottfried 6
limit of a vector function 183
limiting friction 132
line of action of a force 143
linear differential equation 16, 34
logistic equation 6, 16, 20

magnitude
  of a scalar 97
  of a vector 97, 98, 109
mass 124, 193, 194
maximum height 218, 220
maximum range 218, 226, 227
method of undetermined coefficients 49, 53
  exceptional cases 57
model pivot 142
model pulley 135
model rod 142
model string 127
modelling static friction 132
modulus
  of a real number 97
  of a vector 97

Newton, Isaac 6, 95
newton 124, 193
Newton's first law 95, 194
Newton's laws of motion 95
Newton's second law 95, 193, 194, 197
Newton's third law 95
non-homogeneous differential equation 16, 34
normal reaction force 126

order of a differential equation 7
output 4

parallelepiped 122
parallelogram rule 102
parametric form of a straight line 110
particle 123, 194
particular integral 46, 53
particular solution of a differential equation 8, 33, 59
perpendicular vectors 112
pivot 142
point of action of a force 124
polynomial function 49
population model 4
position 182
position vector 107
principle of superposition 36
procedure
  applying Newton's second law 197
  direct integration 11
  Euler's method 27
  exceptional cases in the method of undetermined coefficients 57
  integrating factor method 19
  method of undetermined coefficients 53
  resolving a vector into components 117
  separation of variables 15
  solving homogeneous linear constant-coefficient second-order differential equations 43
  solving inhomogeneous linear constant-coefficient second-order differential equations 47
  solving statics problems for particles 129
projectile 178
  achievable target 223
  equation of motion 216, 220
  impact 215
  in flight 215
  launch 215
  launch angle 216
  launch velocity 215
  maximum height 218, 220
  maximum range 218, 226, 227
  range 217, 221, 225
  time of flight 215
  trajectory 215, 221, 223
proportionate birth rate 4
proportionate death rate 4
proportionate growth rate 5
pulley 135

range 217, 221, 225
reaction force 146
resolving a vector into components 116, 117
resultant vector 101
right-hand grip rule 106

right-hand rule 106
right-handed system 106
rigid body 141
rod 142
rounding error 31

scalar 95, 97
scalar multiple 99, 100
scalar multiplication 99, 100
scalar product 111
scalar triple product 122
scaling a vector 99, 100, 103
  in component form 108
separable differential equation 15
separation of variables method 15
sinusoidal function 52
slipping 130, 138
slope 21, 24
smooth surface 132
solution of a differential equation 7
solving homogeneous equations 43
solving inhomogeneous equations 47
speed 97, 183
static friction 132
statics 96
step length 26
step size 26, 29
string 127
subtraction of vectors 102
sum of vectors 101
  in component form 108

tension force 127
tension in a string 127
terminal speed 212, 214
torque 143
trajectory 215, 221, 223
transpose symbol 107
trial solution 49
triangle rule 101

undetermined coefficients, method of 49, 53
  exceptional cases 57
unit vector 100, 107

vector 95, 97
vector addition 101, 103
  in component form 108
vector addition rule 101
vector equation of a straight line 110
vector function 179
vector product 118
vector subtraction 102
velocity 97, 98, 183
volume of a parallelepiped 122

weight 124, 125

(x, y)-plane 105

zero vector 98, 102, 103